

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,010
FILING DATE: 21-May-2002
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 09/091,873
FILING DATE: 25-JUN-1998
APPLICATION NUMBER: PCT/GB96/03164
FILING DATE: 19-DEC-1996
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26
Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486
RESULT 11
US-09-833-041-18
Sequence 18, Application US/09833041
Publication No. US20030125247A1
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A.
APPLICANT: Hasseltine, William A.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF545
CURRENT FILING DATE: 2001-04-12
PRIORITY APPLICATION NUMBER: US/09/833,041
PRIORITY FILING DATE: 2001-04-12
PRIORITY APPLICATION NUMBER: 60/229,358
PRIORITY FILING DATE: 2000-04-12
PRIORITY APPLICATION NUMBER: 60/256,931
PRIORITY FILING DATE: 2000-12-21
PRIORITY APPLICATION NUMBER: 60/199,384
PRIORITY FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-833-041-18
Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486
RESULT 12
US-09-833-041-18
Sequence 18, Application US/09833041
Publication No. US20030125247A1
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A.
APPLICANT: Hasseltine, William A.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF545
CURRENT FILING DATE: 2001-04-12
PRIORITY APPLICATION NUMBER: US/09/833,041
PRIORITY FILING DATE: 2001-04-12
PRIORITY APPLICATION NUMBER: 60/229,358
PRIORITY FILING DATE: 2000-04-12
PRIORITY APPLICATION NUMBER: 60/256,931
PRIORITY FILING DATE: 2000-12-21
PRIORITY APPLICATION NUMBER: 60/199,384
PRIORITY FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-833-041-18
Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486
RESULT 13
US-09-932-322-445
Sequence 445, Application US/09932322
Publication No. US20030194743A1
GENERAL INFORMATION:
APPLICANT: Dyax Corp.
APPLICANT: Beltzer, James P.
APPLICANT: Potter, M. Daniel
APPLICANT: Fleming, Tony J.
APPLICANT: Ladner, Robert Charles
TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (Blys)
FILE REFERENCE: DYX-018.1 PCT: DYX-018.1 US
CURRENT FILING DATE: 2001-08-17
PRIORITY APPLICATION NUMBER: US/09/932,322
PRIORITY FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 458
SOFTWARE: PatentIn version 3.1
SEQ ID NO 445
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-932-322-445
Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486
RESULT 14
US-09-832-501-18
Sequence 18, Application US/09832501
Publication No. US20030199043A1
GENERAL INFORMATION:
APPLICANT: Ballance, David J.
APPLICANT: Sleep, Darrell

US-09-833-117-18
Sequence 18, Application US/09833117
Publication No. US20030171267A1
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A.
APPLICANT: Sadeghi, Homa
APPLICANT: Prior, Christopher P.
APPLICANT: Turner, Andrew J.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF543
CURRENT APPLICATION NUMBER: US/09/833,117
CURRENT FILING DATE: 2001-04-12
PRIORITY APPLICATION NUMBER: 60/229,358
PRIORITY FILING DATE: 2000-04-12
PRIORITY APPLICATION NUMBER: 60/256,931
PRIORITY FILING DATE: 2000-12-21
PRIORITY APPLICATION NUMBER: 60/199,384
PRIORITY FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-833-117-18
Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486
RESULT 13
US-09-932-322-445
Sequence 445, Application US/09932322
Publication No. US20030194743A1
GENERAL INFORMATION:
APPLICANT: Dyax Corp.
APPLICANT: Beltzer, James P.
APPLICANT: Potter, M. Daniel
APPLICANT: Fleming, Tony J.
APPLICANT: Ladner, Robert Charles
TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (Blys)
FILE REFERENCE: DYX-018.1 PCT: DYX-018.1 US
CURRENT FILING DATE: 2001-08-17
PRIORITY APPLICATION NUMBER: US/09/932,322
PRIORITY FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 458
SOFTWARE: PatentIn version 3.1
SEQ ID NO 445
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-932-322-445
Query Match 100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TESLVNRRP 9
DB 478 TESLVNRRP 486
RESULT 14
US-09-832-501-18
Sequence 18, Application US/09832501
Publication No. US20030199043A1
GENERAL INFORMATION:
APPLICANT: Ballance, David J.
APPLICANT: Sleep, Darrell

Mon Apr 19 13:27:24 2004

```

; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 45; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TESLVNRRP 9
DB      478 TESLVNRRP 486

RESULT 15
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 45; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TESLVNRRP 9
DB      478 TESLVNRRP 486

Search completed: April 19, 2004, 12:55:00
Job time : 4.78947 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.3795 Seconds
(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_439_447

Perfect score: 51

Sequence: 1 KHPEAKRMP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	585	1	US-08-153-799-14
2	51	100.0	585	1	US-08-448-196A-3
3	51	100.0	585	2	US-08-984-176-1
4	51	100.0	585	3	US-08-702-572-2
5	51	100.0	585	3	US-08-769-746-2
6	51	100.0	585	4	US-10-153-064-5
7	51	100.0	609	1	US-08-222-819-3
8	51	100.0	609	1	US-08-433-037-4
9	51	100.0	609	4	US-08-897-956A-2
10	51	100.0	609	4	US-10-153-064-7
11	51	100.0	609	4	US-09-976-594-977
12	51	100.0	609	5	PCT-US95-04075-3
13	51	100.0	610	2	US-08-797-689-2
14	51	100.0	610	4	US-08-984-186-2
15	51	100.0	651	4	US-10-153-064-133
16	51	100.0	652	4	US-10-153-064-96
17	51	100.0	652	4	US-10-153-064-99
18	51	100.0	652	4	US-10-153-064-105
19	51	100.0	652	4	US-10-153-064-132
20	51	100.0	653	4	US-10-153-064-131
21	51	100.0	656	4	US-10-153-064-130
22	51	100.0	660	4	US-10-153-064-90
23	51	100.0	660	4	US-10-153-064-93
24	51	100.0	668	4	US-10-153-064-102
25	51	100.0	676	4	US-10-153-064-95
26	51	100.0	676	4	US-10-153-064-98
27	51	100.0	676	4	US-10-153-064-104

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5768883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

ALIGNMENTS

28 51 100.0 676 4 US-10-153-064-127 Sequence 127, App
29 51 100.0 676 4 US-10-153-064-129 Sequence 129, App
30 51 100.0 677 4 US-10-153-064-125 Sequence 125, App
31 51 100.0 680 4 US-10-153-064-123 Sequence 123, App
32 51 100.0 684 4 US-10-153-064-92 Sequence 92, App
33 51 100.0 692 4 US-10-153-064-101 Sequence 101, App
34 51 100.0 783 1 US-08-256-938-2 Sequence 2, Appli
35 51 100.0 787 1 US-08-256-938-4 Sequence 4, Appli
36 51 100.0 787 1 US-08-797-689-16 Sequence 16, Appl
37 51 100.0 787 4 US-09-984-186-16 Sequence 16, Appl
38 51 100.0 787 4 US-08-897-956A-3 Sequence 3, Appli
39 51 100.0 1184 4 US-10-153-064-89 Sequence 89, Appl
40 39 76.5 393 4 US-09-232-991A-19791 Sequence 15791, A
41 37 72.5 474 3 US-08-978-741-8 Sequence 8, Appli
42 37 72.5 474 3 US-08-978-741-8 Sequence 8, Appli
43 36 70.6 268 4 US-09-328-352-8143 Sequence 12, Appl
44 35 68.6 288 3 US-09-335-409-18 Sequence 18, Appl
45 35 68.6 288 3 US-09-335-409-19 Sequence 19, Appl

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: Region
LOCATION: 369..419
OTHER INFORMATION: /note= "Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"
FEATURE:

NAME/KEY: Region
LOCATION: 1..585
OTHER INFORMATION: /note= "Amino acid sequence of
OTHER INFORMATION: natural HSA"
US-08-153-799-14

Query Match 100.0%; Score 51; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
|||||
DB 439 KHPEAKRMP 447

RESULT 2

US-08-448-196A-3
Sequence 3, Application US/08448196A
Patent No. 5780594
GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C.
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
TITLE OF INVENTION: RELATED PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448.196A
FILING DATE: 23-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 585 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

US-08-448-196A-3

Query Match 100.0%; Score 51; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
|||||
DB 439 KHPEAKRMP 447

RESULT 3

US-08-984-176-1
Sequence 1, Application US/08984176

Patent No. 5948609

GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C

APPLICANT: HO, JOSEPH X

APPLICANT: RUKER, FLORIAN

TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984,176

CURRENT APPLICATION NUMBER: US/08/984,176

CURRENT FILING DATE: 1997-12-03

NUMBER OF SEQ ID NOS: 1

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 1

LENGTH: 585

TYPE: PRT

ORGANISM: Homo sapiens

US-08-984-176-1

Query Match 100.0%; Score 51; DB 2; Length 585;

Best Local Similarity 100.0%; Pred. No. 0.065;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
|||||
DB 439 KHPEAKRMP 447

RESULT 4

US-08-702-572-2

Sequence 2, Application US/08702572

Patent No. 5965386

GENERAL INFORMATION:

APPLICANT: Kerry-Williams, Sean M

APPLICANT: Gilbert, Sarah C

TITLE OF INVENTION: Yeast Strains and Modified Albumins

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Centeon L.L.C.

STREET: 1020 First Avenue

CITY: King of Prussia

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19406-1310

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: Microsoft Word 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/702,572

FILING DATE: 11-NOV-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO 95/23857

FILING DATE: 1-MAR-1995

APPLICATION NUMBER: GB 9404270.2

FILING DATE: 5-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Naomi Biswas

REGISTRATION NUMBER: 38,384

REFERENCE/DOCKET NUMBER: CE0114 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610/878/4294

TELEFAX: 610/878/4221

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 51; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9
Db 439 KHPKAKRMP 447

RESULT 5

US-08-769-746-2
Sequence 2, Application US/08769746
Patent No. 6274305
GENERAL INFORMATION:
APPLICANT: Sonnenschein, Carlos
APPLICANT: Soto, Ana M.
TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,746
FILING DATE: 19-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cartoll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: MBRI-02584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-769-746-2

Query Match 100.0%; Score 51; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9
Db 439 KHPKAKRMP 447

RESULT 6

US-10-153-064-5
Sequence 5, Application US/10153064
Patent No. 6663485
GENERAL INFORMATION:
APPLICANT: Bell et al.
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
FILE REFERENCE: PF556

CURRENT APPLICATION NUMBER: US/10/153,064
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: 60/293,212
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 137
SOFTWARE: Patent In version 3.1
SEQ ID NO 5
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 51; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9
Db 439 KHPKAKRMP 447

RESULT 7

US-08-222-619-3
Sequence 3, Application US/08222619
Patent No. 5652352
GENERAL INFORMATION:
APPLICANT: Lichenstein, Henri
APPLICANT: Lyons, David
APPLICANT: Wurfel, Mark
APPLICANT: Wright, Samuel
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,619
FILING DATE:
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-222-619-3

Query Match 100.0%; Score 51; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.065;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9
Db 463 KHPKAKRMP 471

RESULT 8

US-08-433-037-4
Sequence 4, Application US/08433037
Patent No. 5707828
GENERAL INFORMATION:
APPLICANT: Sreekrishna, Kotikanyadan

APPLICANT: Bair, Kathryn A.
APPLICANT: Brierley, Russell A.
APPLICANT: Thill, Gregory P.
APPLICANT: Tschopp, Juerg F.
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
TITLE OF INVENTION: PICHIA PASTORIS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,037
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DiGiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9108Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 100.0%; Score 51; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 463 KHPEAKRMP 471

RESULT 9
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CFA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 100.0%; Score 51; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 463 KHPEAKRMP 471

RESULT 10
US-10-153-064-7
; Sequence 7, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match 100.0%; Score 51; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 463 KHPEAKRMP 471

RESULT 11
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match 100.0%; Score 51; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 463 KHPEAKRMP 471

RESULT 12
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.

PRIOR APPLICATION DATA:

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match 100.0%; Score 51; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. NO. 0.069;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
Db 463 KHPEAKRMP 471

RESULT 15
US-10-153-064-133
; Sequence 133; Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match 100.0%; Score 51; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. NO. 0.074;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
Db 505 KHPEAKRMP 513

Search completed: April 19, 2004, 12:05:21
Job time : 2.3795 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.70637 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: us-09-832-929-18_COPY_462_475

Perfect score: 71

Sequence: 1 VLHEKTPVSDRVTK 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*

1: piri.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	100.0	609	1 ABHUS	serum albumin prec
2	67	94.4	608	2 S57632	serum albumin prec
3	64	90.1	600	2 A47391	serum albumin prec
4	64	90.1	605	1 ABPGS	serum albumin prec
5	64	90.1	607	1 ABBOS	serum albumin prec
6	64	90.1	607	1 ABSHS	serum albumin prec
7	64	90.1	608	1 ABRHS	serum albumin prec
8	63	88.7	607	1 ABHOS	serum albumin prec
9	60	84.5	609	2 UC5838	albumin - Mongolia
10	59	83.1	453	2 A05139	serum albumin - mo
11	43	60.6	553	2 D83640	hypothetical prote
12	43	60.6	733	2 E71937	flagellar biosynth
13	43	60.6	733	2 A64650	flagellar biosynth
14	40	56.3	342	2 S63404	hypothetical prote
15	40	56.3	604	2 B9C289	conserved hypochet
16	40	56.3	602	2 T29784	hypothetical prote
17	39	54.9	577	2 T45333	hypothetical prote
18	39	54.9	628	2 B84415	translation elonga
19	39	54.9	728	1 S07558	probable DNA-bind
20	39	54.9	753	2 G87178	methane monooxygen
21	38	53.5	169	2 C39049	hypothetical prote
22	38	53.5	172	2 F70789	lipase lipC PA4813
23	38	53.5	309	2 G83044	serine-tRNA ligase
24	38	53.5	457	2 T02786	hypothetical prote
25	38	53.5	539	2 T02705	hypothetical prote
26	38	53.5	690	2 F97164	flagellar biosynth
27	37	52.1	277	2 T36312	hypothetical prote
28	37	52.1	257	2 AB3642	protein ybis precu
29	37	52.1	383	2 T50775	probable translati

hemolin - cecropia
hemolin precursor
neutral metallopro
hypothetical prote
hypothetical prote
hypothetical Myb f
pled protein - Syn
probable translati
translation initia
probable translati
phosphoribosylamin
filamentous hemagg
hypothetical prote
hypothetical prote
lmbN protein - Str
DNA-directed DNA p

30 37 52.1 413 2 S65948
31 37 52.1 413 2 A37778
32 37 52.1 438 2 S42581
33 37 52.1 477 2 T16546
34 37 52.1 482 2 G96616
35 37 52.1 661 2 T38176
36 37 52.1 829 2 S75776
37 37 52.1 900 2 T47732
38 37 52.1 900 2 T50773
39 37 52.1 935 2 T50774
40 37 52.1 1353 1 AUFFEM
41 37 52.1 4152 2 T31102
42 37 52.1 4919 2 T31105
43 36 50.7 113 2 A81439
44 36 50.7 275 2 S44961
45 36 50.7 331 2 T09748

ALIGNMENTS

RESULT 1

ASHUS
serum albumin precursor [validated] - human
N:Alternate names: preproalbumin
N:Contains: kinetensin
C:Species: Homo sapiens (man)
C:Dates: 29-Jul-1981 #sequence revision 31-Jan-1997 #text-change 17-Mar-2000
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S31
R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu:
Nucleic Acids Res. 9, 6103-6114, 1981
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia col:
A:Reference number: A93743; MUID:82081882; PMID:6171778
A:Accession: A93743
A:Molecule type: mRNA
A:Residues: 1-419, 'K', 421-609 <LAW>
A:Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA2:
R:Dugaiczyk, A.; Law, S.W.; Dennison, O.E
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A:Reference number: A93936; MUID:82105994; PMID:6275391
A:Accession: A93936
A:Molecule type: mRNA
A:Residues: 1-120, 'G', 122-609 <DUG>
A:Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590
R:Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoka, T.
J. Biol. Chem. 261, 3244-3251, 1986
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A:Reference number: I39427; MUID:86140099; PMID:2419329
A:Accession: I39427
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-26 <URA>
A:Cross-references: GB:IMJ3075; NID:G178330; PIDN:AAA51688.1; PID:G553173
R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.
A:Reference number: I59286; MUID:94181575; PMID:8134387
A:Accession: I59286
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 282-290, 'KSRPDLQ' <WAT>
A:Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033
A>Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-
A:Reference number: I59313; MUID:94294404; PMID:8022807
A:Accession: I59313
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 589-590, 'ALPRVKNLLQVLP' <MAD>
A:Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; PID:G547232

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants
R;Menaya, J.; Parvillia, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A;Reference number: G08292
A;Accession: G01747
A;Status: translated from GS/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-120, 'G', 122-455 <MEN>
A;Cross-references: EMBL:U22961; NID:q763428; PID:AAA64922.1; PID:g763431
R;Ledgerwood, B.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
A;Reference number: S55314; MUID:95275251; PMID:7755581
A;Accession: S55314
A;Molecule type: Protein
A;Residues: 19-27 <LED>
R;Meloun, B.; Moravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A;Title: Complete amino acid sequence of human serum albumin.
A;Reference number: A91420; MUID:76187907; PMID:1225573
A;Accession: A91420
A;Molecule type: Protein
A;Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-
R;Roehr, U.; Spittler, G.; Tripiet, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from un
A;Reference number: S06422
A;Note: this paper is in German, with an English abstract
A;Accession: S06422
A;Molecule type: Protein
A;Residues: 25-48 <RO>
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 593-599, 1993
A;Title: Mass spectrometric identification of modifications to human serum albumin treat
A;Reference number: S36882; MUID:93384321; PMID:8373198
A;Accession: S36882
A;Molecule type: Protein
A;Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <FIN>
R;Kausler, E.; Spittler, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A;Title: Bruchstuecke aus Albumin und Beta(2)-Mikroglobulin - Bestandteile der Mittelmol
A;Reference number: S17599; MUID:92126241; PMID:1772598
A;Accession: S17599
A;Molecule type: Protein
A;Residues: 25-54, 354-357, 431-447 <KAU>
A;Note: 49-Leu was also found
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: A45800
A;Molecule type: Protein
A;Residues: 166-173 <CAR>
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre
A;Reference number: A03239; MUID:86242180; PMID:3087352
A;Accession: A03239
A;Molecule type: Protein
A;Residues: 166-173, 'L', <MOG>
R;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A;Title: Mutations in genetic variants of human serum albumin found in Italy.
A;Reference number: A38255; MUID:91062352; PMID:2247440
A;Accession: A38255
A;Molecule type: Protein
A;Residues: 76-111 <GAL1>
A;Accession: B38255
A;Molecule type: Protein
A;Residues: 82-105, 'K', 107-110 <GAL2>
A;Note: this variant is designated albumin Vibo Valentia
A;Accession: A38255
A;Molecule type: Protein

A;Residues: 76-83, 'K', 85-106 <GAL3>
A;Note: this variant is designated albumin Torino
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin He
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: Protein
A;Residues: 255-263, 'E', 265-281 <MIN1>
A;Note: this variant is designated albumin Herborn
R;Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,
Biochim. Biophys. Acta 1119, 233-238, 1992
A;Title: Two albumins with identical electrophoretic mobility are produced by diffe
A;Reference number: S21078; MUID:92190239; PMID:1347703
A;Accession: S21078
A;Molecule type: Protein
A;Residues: 354-356, 'K', 358-378 <MIN2>
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported.
R;He, X.M.; Carter, D.C.
Nature 359, 209-215, 1992
A;Title: Atomic structure and chemistry of human serum albumin.
A;Reference number: A46756; MUID:92334427; PMID:1630489
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40
A;Reference number: A94442
A;Contents: annotation, three-dimensional structure and disulfide bonds
R;Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A;Title: Disulfide bonds in human serum albumin.
A;Reference number: A90930
A;Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A;Reference number: A90299; MUID:78186830; PMID:656055
A;Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjoelholm, I., eds., 11-20,
A;Title: Serum albumin: conformation and active sites.
A;Reference number: A94408
A;Contents: annotation; binding sites
R;Harper, M.E.; Dugaiczky, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene:
A;Reference number: A90028; MUID:83279982; PMID:6192711
A;Contents: annotation; gene position
R;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid
A;Reference number: A46755; MUID:76257808; PMID:955075
A;Contents: annotation
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic aci
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosph
A;Reference number: A56294; MUID:92183881; PMID:1544460
A;Contents: annotation
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in l
ataze activity
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
C;Comment: A large number of variants of human serum albumin have been described.
C;Genetics:
A;Gene: GDB:ALB
A;Cross-references: GDB:118990; OMIM:103600
A;Map position: 4q11-4q13
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyrido
F.118/Domain: signal sequence #status predicted <SIG>
F.125-24/Domain: propeptide #status experimental <PRO>
F.125-609/Product: serum albumin #status experimental <MPT>
F.129-202/Domain: serum albumin repeat homology <SAL>

F:166-174/Product: kinetensin #status experimental <KIP>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:217/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4
F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 71; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 7.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
|||||:|||||
Db 486 VLHEKTPVSDRVTK 499

RESULT 2

serum albumin precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
C:Accession: J04660; S57632
R:Hilger, C.; Grigioni, F.; Hentges, F.

Gene 169, 295-296, 1996
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A:Reference number: J04660; MUID:96194824; PMID:8647469

A:Accession: J04660
A:Molecule type: mRNA

A:Residues: 1-608 <H12>
A:Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485

A:Experimental source: liver
C:Comment: This protein is the major protein component in plasma. It functions as a multi-

chain has 35 conserved cysteine residues.
C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: liver; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRP>
F:25-608/Product: serum albumin #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match 94.4%; Score 67; DB 2; Length 608;
Best Local Similarity 92.9%; Pred. No. 0.00042;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
|||||:|||||
Db 486 VLHEKTPVSDRVTK 499

RESULT 3

serum albumin precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A47391

R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993

A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin

A:Reference number: A47391; MUID:93211971; PMID:8460152

A:Contents: B/B homozygote

A:Accession: A47391

A:Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-600 <MAT>

A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295

A:Experimental source: liver

A:Note: sequence extracted from NCBI backbone (NCBI:128280, NCBI:128281)

C:Superfamily: serum albumin; serum albumin repeat homology

F:21-194/Domain: serum albumin repeat homology <SA1>

F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 90.1%; Score 64; DB 2; Length 600;
Best Local Similarity 85.7%; Pred. No. 0.0015;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
|||||:|||||
Db 478 VLHEKTPVSEKVTK 491

RESULT 4

ABPGS

serum albumin precursor - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999

C:Accession: S01382; A61006

R>Weinstock, J.; Baldwin, G.S.

Nucleic Acids Res. 16, 9045, 1988

A:Title: Nucleotide sequence of porcine liver albumin.

A:Reference number: S01382; MUID:89016582; PMID:3174440

A:Accession: S01382

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-605 <WEI>

A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798

R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.

J. Bone Miner. Res. 4, 235-241, 1989

A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of minera

A:Reference number: A61006; MUID:89269769; PMID:2728927

A:Accession: A61006

A:Molecule type: protein

A:Residues: 23-51, 'X', '53-54', 'XXXGY', '146', 'E', '148', 'E', '150-151', 'XVN', '155' <LIM>

A:Experimental source: dental enamel

A:Note: albumin and other serum proteins are also found in bone

C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, ceroid hormones (weak bonds with these hormones promote their transfer across the membr.

C:Superfamily: serum albumin; serum albumin repeat homology

F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>

F:17-22/Domain: propeptide #status predicted <PRO>

F:23-605/Product: serum albumin #status predicted <MAT>

F:27-199/Domain: serum albumin repeat homology <SA1>

F:218-391/Domain: serum albumin repeat homology <SA2>

F:410-589/Domain: serum albumin repeat homology <SA3>

F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,

F:261/Binding site: bilirubin (Lys) #status predicted

Query Match 90.1%; Score 64; DB 1; Length 605;
Best Local Similarity 85.7%; Pred. No. 0.0015;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
|||||:|||||
Db 483 VLHEKTPVSEKVTK 496

RESULT 5

ABBS

serum albumin precursor [validated] - bovine

N:Alternate names: 67K protein; preproalbumin

C:Species: Bos primigenius taurus (cattle)

C:Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000

C:Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A9

R:Holowachuk, E.W.; Stoltenberg, J.K.; Reed, R.G.; Peters Jr., T.

submitted to the EMBL Data Library, August 1991

A:Description: Bovine serum albumin: cDNA sequence and expression.

A:Reference number: A38885

A:Accession: A38885

A:Molecule type: mRNA

A:Residues: 1-607 <HOL>

C:Superfamily: serum albumin; serum albumin repeat homology

F:21-194/Domain: serum albumin repeat homology <SA1>

F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>

R:Hiroyama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.

Biochem. Biophys. Res. Commun. 173, 639-646, 1990

C;Title: Rapid confirmation and revision of the primary structure of bovine serum albumin
A;Reference number: A36401; MUID:91083649; PMID:260975
A;Accession: A36401
A;Molecule type: protein
A;Residues: 25-41, 'H', 43-189, 'B', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HR>
R;MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-485, 1979
A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A;Reference number: A91258; MUID:80024278; PMID:488109
A;Accession: A91258
A;Molecule type: protein
A;Residues: 1-32 <MAG>
R;Haieh, J.C.; Lin, F.P.; Tam, M.F.
Anal. Biochem. 170, 1-8, 1988
A;Title: Electrophoretic onto glass-fiber filter from an analytical isoelectrofocusing gel
A;Reference number: A60808; MUID:88267456; PMID:3389500
A;Accession: B60808
A;Molecule type: protein
A;Residues: 25-41 <HSI>
R;Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelins' is albumin
A;Reference number: S10780; MUID:90336641; PMID:2379503
A;Accession: S10780
A;Molecule type: protein
A;Residues: 25-41, 'H', 43-57, 59-64 <STR>
R;Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteases
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: D45800
A;Molecule type: protein
A;Residues: 163-172 <CAR>
R;Caraway, R.E.; Mitra, S.P.; Cochran, D.E.
J. Biol. Chem. 262, 5968-5973, 1987
A;Title: Structure of a biologically active neurotensin-related peptide obtained from peptidase
A;Reference number: A26693; MUID:87194805; PMID:2437111
A;Accession: A26693
A;Molecule type: protein
A;Residues: 163-172, 'L', <CA2>
R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.
Biochem. J. 191, 867-868, 1980
A;Title: Sequence of residues 400-403 of bovine serum albumin.
A;Reference number: A90309; MUID:82023364; PMID:7283978
A;Accession: A90309
A;Molecule type: protein
A;Residues: 402-433 <REE>
R;Brown, J.R.
Fed. Proc. 34, 591, 1975
A;Title: Structure of bovine serum albumin.
A;Reference number: A91458
A;Accession: A91458
A;Molecule type: protein
A;Residues: 25-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'B', 191-194, 'A', 196-213, 'T', 215-288, 'E'
R;Brown, J.R.
submitted to the Atlas, April 1975
A;Reference number: A94551
A;Accession: A94551
A;Molecule type: protein
A;Residues: 190-195 <BR2>
R;Brown, J.R.
Fed. Proc. 33, 1389, 1974
A;Reference number: A91457
A;Contents: annotation; disulfide bonds
R;Werlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC 3.4.21.3)
A;Reference number: S55232; MUID:95031935; PMID:7945219
A;Accession: S55232
A;Status: preliminary
A;Molecule type: protein
A;Residues: 529-536; 569-572 <WER>
C;Superfamily: serum albumin; serum albumin repeat homology

C;Keywords: carrier protein; copper binding; duplication; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-607/Product: serum albumin #status experimental <MPT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,
Query Match 90.1%; Score 64; DB 1; Length 607;
Best Local Similarity 85.7%; Pred. NO. 0.0015;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLHEKTPVSDRVTK 14
DB 485 VLHEKTPVSEKVK 498
RESULT 6
ABSHS
serum albumin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C;Accession: S06936
R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A;Reference number: S06936; MUID:90098888; PMID:2602160
A;Accession: S06936
A;Molecule type: mRNA
A;Residues: 1-607 <BRO>
A;Cross-references: EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membr
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,
Query Match 90.1%; Score 64; DB 1; Length 607;
Best Local Similarity 85.7%; Pred. NO. 0.0015;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLHEKTPVSDRVTK 14
DB 485 VLHEKTPVSEKVK 498
RESULT 7
ABRTS
serum albumin precursor - rat
N;Alternate names: preproalbumin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 22-Jun-1999
C;Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R;Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A;Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A;Reference number: A93872; MUID:81223722; PMID:7017712
A;Accession: A93872
A;Molecule type: mRNA
A;Residues: 1-608 <SAR>
A;Cross-references: GB:V01222; GB:J00698; NID:955627; PIDN:CAA24532.1; PID:955628
R;Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977

A;Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
A;Reference number: A92211; MUID:77249657; PMID:893447
A;Note: Cleavages during protein maturation
A;Accession: A92211
A;Molecule type: protein
A;Residues: 1-38 <STR>
R;Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A;Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A;Reference number: A91946; MUID:78109429; PMID:564345
A;Accession: A91946
A;Molecule type: protein
A;Residues: 25-222 <TS1>
R;Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A;Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid
A;Reference number: A91940; MUID:76260153; PMID:956149
A;Accession: A91940
A;Molecule type: protein
A;Residues: 223-288; 572-608 <IS2>
A;Note: 262-Leu was also found
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 36, 3483-3486, 1976
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.
A;Reference number: A90758; MUID:79001617; PMID:80265
A;Contents: annotation; copper binding
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteases
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: A45800
A;Status: preliminary
A;Molecule type: protein
A;Residues: 166-173 <CAR>
R;Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A;Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved
A;Reference number: I57621; MUID:87286876; PMID:3475566
A;Accession: I57621
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: GB:M16825; NID:G202828; PIDN:AAA40712.1; PID:G554412
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-608/Product: serum albumin #status experimental <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status experimental
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4

Query Match 90.1%; Score 64; DB 1; Length 608;
Best Local Similarity 85.7%; Pred. No. 0.0015;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 486 VLHEKTPVSEKVTK 499
|||||:|||||

RESULT 8
ASHOS
serum albumin precursor - horse
C;Species: Equus caballus (domestic horse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
A;Accession: S34053
R;Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A;Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A;Reference number: S34053; MUID:93345495; PMID:834282

A;Accession: S34053
A;Molecule type: mRNA
A;Residues: 1-607 <HOA>
A;Cross-references: GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membr.
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,
F;263/Binding site: bilirubin (Lys) #status predicted

Query Match 88.7%; Score 63; DB 1; Length 607;
Best Local Similarity 78.6%; Pred. No. 0.0023;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 485 VLHEKTPVSEKVTK 498
|||||:|||||

RESULT 9
JCS838
albumin - Mongolian jird
C;Species: Meriones unguiculatus (Mongolian jird)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000
A;Accession: JCS838
R;Yoshida, K.; Seto-Oshima, A.; Sinohara, H.
DNA Res. 4, 351-354, 1997
A;Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in th
A;Reference number: JCS838; MUID:98116663; PMID:9455485
A;Accession: JCS838
A;Molecule type: mRNA
A;Residues: 1-609 <YOS>
A;Cross-references: DBJ:AB006197; NID:G2317277; PIDN:BAA21765.1; PID:G2317278
A;Experimental source: liver
C;Superfamily: serum albumin; serum albumin repeat homology
F;222-395/Domain: serum albumin repeat homology <SA2>

Query Match 84.5%; Score 60; DB 2; Length 609;
Best Local Similarity 78.6%; Pred. No. 0.008;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 487 LLHEKTPVSEQVTK 500
|||||:|||||

RESULT 10
A05139
serum albumin - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
A;Accession: A05139; I48638
R;Minghetti, P.P.; Law, S.W.; Dugaiczky, A.
Mol. Biol. Evol. 2, 347-358, 1985
A;Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudo;
A;Reference number: A93055; MUID:88216123; PMID:2452956
A;Accession: A05139
A;Molecule type: mRNA
A;Residues: 1-418 <MIN>
A;Cross-references: GB:M16111; NID:G191764; PIDN:AAA37190.1; PID:G191765
R;Boccaccio, C.; Deschattrette, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A;Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in th
A;Reference number: I48638; MUID:90269606; PMID:1971802
A;Accession: I48638

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 379-453 <BOC>
A;Cross-references: EMBL:U13060; NID:G52939; PIDN:CAA31458.1; PID:G899334
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; Plasma
F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F;123-296/Domain: serum albumin repeat homology <SA2>
F;315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match 83.1%; Score 59; DB 2; Length 453;
Best Local Similarity 78.6%; Pred. No. 0.0088;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
:|||||:|
Db 388 LLHEKTPVSEHVK 401

RESULT 11
D83640
Hypothetical protein PA0049 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: D83640
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:2043737; PMID:10984043
A;Accession: D83640
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-553 <STO>
A;Cross-references: GB:AE004444; GB:AE004091; NID:G9945863; PIDN:AG034339.1; GSPDB:GN0014
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0049

Query Match 60.6%; Score 43; DB 2; Length 553;
Best Local Similarity 42.9%; Pred. No. 9.2;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
:|||||:|
Db 71 VTHDQTPLEAERVK 84

RESULT 12
E71937
flagellar biosynthesis protein - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 24-Sep-1999
C;Accession: E71937
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: E71937
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-733 <ARN>
A;Cross-references: GB:AE001473; GB:AE001439; NID:G4154910; PIDN:AAD05964.1; PID:G4154910
A;Experimental source: strain J99
C;Genetics:
A;Gene: flaB
C;Superfamily: regulatory protein lcrD

Query Match 60.6%; Score 43; DB 2; Length 733;
Best Local Similarity 53.8%; Pred. No. 13;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 13
:|||||:|
Db 577 LLHEKIPKIDMLT 589

RESULT 13
A64650
flagellar biosynthesis protein - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 24-Sep-1999
C;Accession: A64650
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 533-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: A64650
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-733 <ROM>
A;Cross-references: GB:AE000612; GB:AE000511; NID:G2314188; PIDN:AAD08087.1; PID:G2314188
C;Superfamily: regulatory protein lcrD

Query Match 60.6%; Score 43; DB 2; Length 733;
Best Local Similarity 53.8%; Pred. No. 13;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 13
:|||||:|
Db 577 LLHEKIPKIDMLT 589

RESULT 14
S63404
hypothetical protein YNR071c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein N3605
C;Species: Saccharomyces cerevisiae
C;Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 19-Apr-2002
C;Accession: S63404
R;Andre, B.; Iraqui Housasini, I.; Urrestarazu, L.A.; Vissers, S.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S62920
A;Accession: S63404
A;Molecule type: DNA
A;Residues: 1-342 <AND>
A;Cross-references: EMBL:Z71686; NID:G1302605; PID:G1302606; GSPDB:GN00014; MIPS:YNR071c
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YNR071c
A;Cross-references: SGD:S0005354
A;Map position: 14R
C;Superfamily: aldose 1-epimerase

Query Match 56.3%; Score 40; DB 2; Length 342;
Best Local Similarity 80.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSD 10
:|||||
Db 231 VLHEDTPVFD 240

RESULT 15
B90289
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: B90289
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: B90289
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-604 <KUR>
A;Cross-references: GB:AE006641; NID:gl3814539; PIDN:AAK41569.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO1331

Query Match 56.3%; Score 40; DB 2; Length 604;
Best Local Similarity 61.5%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVT 13
:|||||
Db 6 ILHETTQVKDLVT 18

Search completed: April 19, 2004, 12:02:32
Job time : 2.70637 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 1.02124 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_462_475

Perfect score: 71

Sequence: 1 VLHEKTPVSDRVTK 14

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	100.0	609	1	ALBU HUMAN
2	67	94.4	608	1	ALBU CANFA
3	67	94.4	608	1	ALBU FELCA
4	64	90.1	600	1	ALBU MACWU
5	64	90.1	605	1	ALBU PIG
6	64	90.1	607	1	ALBU BOVIN
7	64	90.1	607	1	ALBU SHEEP
8	64	90.1	608	1	ALBU RABIT
9	64	90.1	608	1	ALBU RAT
10	63	88.7	607	1	ALBU HORSE
11	60	84.5	609	1	ALBU MERUN
12	59	83.1	608	1	ALBU MOUSE
13	43	60.6	733	1	FLHA HELPJ
14	43	60.6	733	1	FLHA HELPY
15	42	59.2	905	1	NIQG PSESM
16	40	56.3	342	1	YN9A YEAST
17	39	54.9	346	1	CATV GVXN
18	39	54.9	728	1	EF2 HALHA
19	38	53.5	169	1	MEMG METTR
20	38	53.5	539	1	FUT2 ARATH
21	37	52.1	413	1	HEMO HYACE
22	37	52.1	438	1	SHPI STAHY
23	37	52.1	505	1	YMR2 CAERL
24	37	52.1	661	1	YBJU SCHEO
25	37	52.1	900	1	IF38 ARATH
26	37	52.1	935	1	IF38 MEDTR
27	37	52.1	1353	1	PUR2 DROME
28	36	50.7	256	1	Y9G8 BRAJA
29	36	50.7	275	1	LMEN STRLN
30	36	50.7	331	1	DP3Z MYOGE
31	36	50.7	514	1	GPW1 SHEON
32	36	50.7	601	1	ZF37 RAT
33	36	50.7	643	1	GYRB MYCCLA

ALIGNMENTS

RESULT 1

ALBU_HUMAN STANDARD; PRT; 609 AA.
 ID AC P02768; O95574; Q13140; Q9P157; Q9UHS3; Q9UJ20;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCB:TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86196112; PubMed=3009475;
 RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
 RA Beattie W.G., Dugaiczky A.;
 RT "Molecular structure of the human albumin gene is revealed by
 RT nucleotide sequence within q11-22 of chromosome 4.";
 RL J. Biol. Chem. 261:6747-6757(1986).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
 RX MEDLINE=82081882; PubMed=6171778;
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
 RA Najarian R.C., Seeburg P.H., Wion K.L.;
 RT "The sequence of human serum albumin cDNA and its expression in E.
 RT coli.";
 RL Nucleic Acids Res. 9:6103-6114(1981).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
 RX MEDLINE=82105994; PubMed=6275391;
 RA Dugaiczky A., Law S.W., Dennison O.E.;
 RT "Nucleotide sequence and the encoded amino acids of human serum
 RT albumin mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).
 RC TISSUE=fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
 Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
 RT "Functional prediction of the coding sequences of 121 new genes
 RT deduced by analysis of cDNA clones from human fetal liver.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
 RA Huang M.C., Wu H.T.;
 RT "The cDNA sequences of human serum albumin.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.

34 36 50.7 734 1 PSAB CVACA O9t06 cyanidium c
 35 36 50.7 824 1 DD20_HUMAN O3uh6 homo sapien
 36 36 50.7 964 1 DPOL_CBEFV P30319 choristoneu
 37 35 49.3 154 1 PHBS_STAEP Q8cn8 staphylococ
 38 35 49.3 183 1 NUPM_NEUCR P21976 neurospora
 39 35 49.3 255 1 AQOD_BACSU P35146 bacillus su
 40 35 49.3 271 1 AQP2_MOUSE P56402 mus musculu
 41 35 49.3 271 1 AQP2_RAT P34080 rattus norv
 42 35 49.3 273 1 MADE_STAAM Q99ax5 staphylococ
 43 35 49.3 321 1 YBAX_ECOLI P76254 escherichia
 44 35 49.3 327 1 SYFA_PASMU P57860 pasteurella
 45 35 49.3 328 1 SDHL_HUMAN P20132 homo sapien

RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Teisgold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RX SEQUENCE OF 25-609.
RP MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RX SEQUENCE OF 25-609.
RP Brown J.R., Shockley P., Behrens P.Q.;
RA (In) Bing D.H. (eds.);
RL The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RX SEQUENCE OF 1-455 FROM N.A.
RP TISSUE=Liver;
RC Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11]
RX SEQUENCE OF 1-26 FROM N.A.
RP MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RX SEQUENCE OF 222-229.
RP MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RX SEQUENCE OF 25-44 AND 480-499.
RP TISSUE=Heart;
RC MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994.";
RL Electrophoresis 15:1459-1465(1994).
RN [14]
RX DISULFIDE BONDS.
RA Sabar M.A., Stockbauer P., Moravek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15]
RX BILIRUBIN-BINDING SITE.
RP MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin.";
RL Biochem. J. 171:453-459(1978).

RN [16]
RX VARIANT CANTERBURY ASN-337.
RA MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
RT domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RX VARIANTS NAG-2 AND NAG-3.
RA MEDLINE=88068523; PubMed=3479777;
RA Takehashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
RT Amerindian and Japanese populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18]
RX VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RA MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese allolalbumins.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19]
RX VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RA MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20]
RX DESCRIPTION OF VARIANT REDHILL.
RA MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
RT human serum albumin whose precursor has an aberrant signal peptidase
RT cleavage site.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21]
RX VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RA MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
RT Italy.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22]
RX VARIANT VENEZIA.
RA MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
RT carboxyl-terminal variants of human serum albumin.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23]
RX VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RA KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RA MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Amaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN [24]
RX VARIANT CASEBOOK ASN-518.
RA MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
RT albumin: albumin Casebook (494 Asp-->Asn).";
RL Biochim. Biophys. Acta 1097:49-54(1991).
RN [25]
RX VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RA MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two allolalbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 71; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHKTPTVSDRVTK 14
 DB 486 VLHKTPTVSDRVTK 499

RESULT 2
 ALBU CANFA STANDARD; PRT; 608 AA.
 AC P49222; Q77705; Q9TS24;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serum albumin precursor (Allergen Can f 3).
 GN ALB.

OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Beagle; TISSUE=Liver;
 RA Hilger C.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=20148667; PubMed=10669848;
 RA Pandjaitan B., Swoboda I., Brandejesky-Pichler F., Rumpold H.,
 VA Valenta R., Spitzauer S.;
 RT "Escherichia coli expression and purification of recombinant dog
 albumin, a cross-reactive animal allergen.";
 RL J. Allergy Clin. Immunol. 105:279-285(2000).
 RN [3]
 RP SEQUENCE OF 25-48.
 RX MEDLINE=75011422; PubMed=4414013;
 RA Dixon J.W., Sarkar B.;
 RT "Isolation, amino acid sequence and copper(II)-binding properties of
 peptide (1-24) of dog serum albumin.";
 RL J. Biol. Chem. 249:5872-5877(1974).
 RN [4]
 RP SEQUENCE OF 25-38.
 RC TISSUE=Heart;
 RX MEDLINE=98163340; PubMed=9504812;
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
 dog heart proteins.";
 RL Electrophoresis 18:2795-2802(1997).
 RN [5]
 RP SEQUENCE OF 215-478 FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=94201492; PubMed=7512102;
 RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,
 RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
 RT "Molecular characterization of dog albumin as a cross-reactive
 allergen.";
 RL J. Allergy Clin. Immunol. 93:614-627(1994).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- ALLERGEN: Causes an allergic reaction in human.
 CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -1- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AJ133489; CAB64867.1; -;
 CC EMBL; Y17737; CAA76841.1; -;
 CC EMBL; S72946; AAB30434.1; -;
 CC HSPF; P02769; I57B.
 CC HSC-2DPAGE; P49822; DOG.
 CC InterPro; IPR000264; Serum_albumin.
 CC Pfam; PF00273; transport_prot; 3.
 CC PRINTS; PR00802; SERUMALBUMIN.
 CC ProDom; PD002486; Serum_albumin; 1.
 CC SMART; SMC0103; ALBUMIN; 3.
 CC PROSITE; PS00212; ALBUMIN; 3.
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
 KW Metal-binding; Lipid-binding; POTENTIAL.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608 SERUM ALBUMIN.
 FT DOMAIN 25 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.
 FT DOMAIN 404 595 ALBUMIN 3.
 FT METAL 27 27 COPPER (BY SIMILARITY).
 FT DISULFID 77 86 BY SIMILARITY.
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 148 193 BY SIMILARITY.
 FT DISULFID 192 201 BY SIMILARITY.
 FT DISULFID 224 270 BY SIMILARITY.
 FT DISULFID 269 277 BY SIMILARITY.
 FT DISULFID 289 303 BY SIMILARITY.
 FT DISULFID 302 313 BY SIMILARITY.
 FT DISULFID 340 385 BY SIMILARITY.
 FT DISULFID 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.
 FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT CONFLICT 1 26 MKVTFISLFLPFSAYSRGLVREA -> MDT (IN
 REF. 2).
 FT CONFLICT 146 146 A -> R (IN REF. 2).
 FT CONFLICT 206 206 I -> T (IN REF. 2).
 FT CONFLICT 349 349 V -> A (IN REF. 2).
 FT CONFLICT 359 359 S -> A (IN REF. 1).
 FT CONFLICT 448 448 V -> VV (IN REF. 5).
 FT CONFLICT 474 474 D -> E (IN REF. 1).
 SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;

Query Match 94.4%; Score 67; DB 1; Length 608;
 Best Local Similarity 92.9%; Pred. No. 0.00035;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHKTPTVSDRVTK 14
 DB 486 VLHKTPTVSDRVTK 499

RESULT 3
 ALBU FELCA STANDARD; PRT; 608 AA.
 ID ALBU FELCA
 AC P49064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serum albumin precursor (Allergen Fel d 2).
 GN ALB.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96194824; PubMed=8647469;
 RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
 RT "Sequence of the gene encoding cat (felis domesticus) serum albumin.";
 RL Gene 169:295-296(1996).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- ALLERGEN: Causes an allergic reaction in human.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC
 CC EMBL; X84842; CAA59279.1; -
 CC PIR; JC4660; S57632.
 CC HSP; P02768; 1E7B.
 CC InterPro; IPR000264; Serum albumin.
 CC Pfam; PF00273; transport prot; 3.
 CC PRINTS; PR00802; SERUMALBUMIN.
 CC ProDom; PD002486; Serum albumin; 1.
 CC SMART; SMO0103; ALBUMIN; 3.
 CC PROSITE; PS00212; ALBUMIN; 3.
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
 KW SIGNAL 1 18
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIN 25 608 SERUM ALBUMIN.
 FT DOMAIN 25 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.
 FT DOMAIN 404 595 ALBUMIN 3.
 FT METAL 27 27 COPPER.
 FT DISULFID 77 86 BY SIMILARITY.
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 148 193 BY SIMILARITY.
 FT DISULFID 132 201 BY SIMILARITY.
 FT DISULFID 224 270 BY SIMILARITY.
 FT DISULFID 269 277 BY SIMILARITY.
 FT DISULFID 289 303 BY SIMILARITY.
 FT DISULFID 302 313 BY SIMILARITY.
 FT DISULFID 340 395 BY SIMILARITY.
 FT DISULFID 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.
 FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 SQ SEQUENCE 608 AA; 68659 MW; 07B629CAC3F60E5F CRC64;
 Query Match 94.4%; Score 67; DB 1; Length 608;
 Best Local Similarity 92.9%; Pred. No. 0.00035;
 Matches 13; Conservative 1; Mismatches 0; Gaps 0;
 OY 1 VLHEKTPVSDRVTK 14
 |||||
 Db 486 VLHEKTPVSDRVTK 499
 RESULT 4
 ALBU_MACMU
 ID _ALBU_MACMU STANDARD; PRT; 600 AA.

AC Q28522;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor (Fragment).
 GN AUB.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93211971; PubMed=8460152;
 RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
 RA Dwyer J., Putnam F.W.;
 RT "cDNA and protein sequence of polymorphic macaque albumins that differ
 RT in bilirubin binding";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC
 CC EMBL; M90463; AAA36906.1; -
 CC PIR; A47391; A47391.
 CC HSP; P02768; 1E7B.
 CC InterPro; IPR000264; Serum albumin.
 CC Pfam; PF00273; transport prot; 3.
 CC PRINTS; PR00802; SERUMALBUMIN.
 CC ProDom; PD002486; Serum albumin; 1.
 CC SMART; SMO0103; ALBUMIN; 3.
 CC PROSITE; PS00212; ALBUMIN; 3.
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 KW NON_TER 1 1
 FT SIGNAL <1 10 BY SIMILARITY.
 FT PROPEP 11 16 BY SIMILARITY.
 FT CHAIN 17 600 SERUM ALBUMIN.
 FT DOMAIN 17 197 ALBUMIN 1.
 FT DOMAIN 204 389 ALBUMIN 2.
 FT DOMAIN 396 587 ALBUMIN 3.
 FT METAL 19 19 COPPER (BY SIMILARITY).
 FT BINDING 256 256 BILIRUBIN (POTENTIAL).
 FT DISULFID 69 78 BY SIMILARITY.
 FT DISULFID 91 107 BY SIMILARITY.
 FT DISULFID 106 117 BY SIMILARITY.
 FT DISULFID 140 185 BY SIMILARITY.
 FT DISULFID 184 193 BY SIMILARITY.
 FT DISULFID 216 262 BY SIMILARITY.
 FT DISULFID 261 269 BY SIMILARITY.
 FT DISULFID 281 295 BY SIMILARITY.
 FT DISULFID 332 377 BY SIMILARITY.
 FT DISULFID 376 385 BY SIMILARITY.
 FT DISULFID 408 454 BY SIMILARITY.
 FT DISULFID 453 464 BY SIMILARITY.
 FT DISULFID 477 493 BY SIMILARITY.
 FT DISULFID 492 503 BY SIMILARITY.
 FT DISULFID 530 575 BY SIMILARITY.
 FT DISULFID 574 583 BY SIMILARITY.
 SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

```
Query Match          90.1%; Score 64; DB 1; Length 600;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
        |||||
        478 VLHEKTPVSEKVTK 491

DB

RESULT 5
ALBU_PIG STANDARD; PRT; 605 AA.
AC P08835; Q29018;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
CX NCBI_TaxID=9823;
EN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S., Weinstein J.;
RT "Nucleotide sequence of porcine liver albumin.";
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X12422; CAA30970.1; -
DR EMBL; M36787; AAA30988.1; -
DR PIR; S01382; ABPGS.
DR HSP; P02768; 1EVH.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON_TER 1
FT PROPEP 41 16 BY SIMILARITY.
FT CHAIN 17 22 BY SIMILARITY.
FT DOMAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.
FT DOMAIN 401 592 ALBUMIN 3.
FT METAL 31 31 COPPER (BY SIMILARITY).
FT METAL 75 84 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.

FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3B556B0DD1A1F4FF CRC64;

Query Match          90.1%; Score 64; DB 1; Length 605;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLHEKTPVSDRVTK 14
        |||||
        483 VLHEKTPVSEKVTK 496

DB

RESULT 6
ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
CX NCBI_TaxID=9913;
EN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX Holowachuk E.W., Stoltenborg J.K., Reed R.G., Peters T. Jr.;
RA Submitted (AUG-1991) to the EMBL/GenBank/DBSJ databases.
RL [2]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RC TISSUE=Liver;
RA Barry T., Power S., Gannon F.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBSJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBSJ databases.
RL [4]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T., Huang M.C.;
RL "The complete cDNA sequence of bovine serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBSJ databases.
RL [5]
RP SEQUENCE OF 1-32. PubMed=488109;
RX MEDLINE=80024278;
RA McGillivray R.T.A., Chung D.W., Davie B.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
RT terminal sequence of prealbumin.";
RL Eur. J. Biochem. 98:477-485(1979).
RL [6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA Brown J.R.;
RT "Structure of bovine serum albumin.";
RL Fed. Proc. 34:591-591(1975).
RL [7]
RP REVISIONS TO 190-195.
RA Brown J.R.;
RL Submitted (APR-1975) to the PIR data bank.
RL [8]
RP SEQUENCE OF 402-433.
```

RX MEDLINE=82023364; PubMed=7283978;
RA Reed R.G., Putnam F.W., Peters T. Jr.;
RL "Sequence of residues 400-403 of bovine serum albumin.";
RN Biochem. J. 191:867-868(1980).
RP SEQUENCE OF 19-28.
RX MEDLINE=77134075; PubMed=843354;
RA Patterson J.E., Geller D.M.;
RL "Bovine microsomal albumin: amino terminal sequence of bovine
proalbumin.";
RN Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.
RX MEDLINE=91083649; PubMed=2260975;
RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;
RL "Rapid confirmation and revision of the primary structure of bovine
serum albumin by EMS and Frit-FAB LC/MS.";
RN Biochem. Biophys. Res. Commun. 173:639-646(1990).
RP SEQUENCE OF 25-41.
RX MEDLINE=98267456; PubMed=3389500;
RA Hsieh J.C., Lin F.P., Tam M.F.;
RL "Electroblotting onto glass-fiber filter from an analytical
isoelectrofocusing gel: a preparative method for isolating proteins
for N-terminal microsequencing.";
RN Anal. Biochem. 170:1-8(1988).
RP SEQUENCE OF 437-451.
RA Vilbois F.;
RL Submitted (AUG-1998) to Swiss-Prot.
RN DISULFIDE BONDS.
RP Brown J.R.;
RL "Structure of serum albumin: disulfide bridges.";
RN Fed. Proc. 33:1389-1393(1974).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- ALLERGEN: Causes an allergic reaction in human.
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL; M73993; AAA51411.1; --
DR EMBL; X59989; CAA4735.1; --
DR EMBL; Y17769; CAA76847.1; --
DR EMBL; AF542068; AAN1824.1; --
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
KW Polymorphism.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).

FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 132
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
FT VARIANT 214 214
FT CONFLICT 302 302
FT CONFLICT 304 305
FT CONFLICT 324 324
FT CONFLICT 394 395
FT CONFLICT 437 437
FT CONFLICT 493 494
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768595D4 CRC64;
Query Match 90.1%; Score 64; DB 1; Length 607;
Best Local Similarity 85.7%; Pred. No. 0.0012; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 2;
QY 1 VLHKTPTVSDRVTK 14
DB 485 VLHKTPTVSEKVTK 498
|||||:|:|:
RESULT 7
ALBU_SHEEP STANDARD; PRT; 607 AA.
AC P14639;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90098888; PubMed=2602160;
RA Brown W.N., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
RL "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
RN Nucleic Acids Res. 17:10495-10495(1989).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

```

DR EMBL; X17055; CAA34903.1; -
DR PIR; S06936; ABSHS.
DR HSP; P02768; IE7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT CHAIN 19 24
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 90.1%; Score 64; DB 1; Length 607;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
Db 485 VLHEKTPVSEKVTK 498

RESULT 8
ALBU RABIT
ID ALBU RABIT STANDARD; PRT; 608 AA.
AC P49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; UL8344; AAB58347.1; -
DR HSP; P02768; IE7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT CHAIN 19 24
FT DOMAIN 25 205
FT DOMAIN 212 397
FT DOMAIN 404 595
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
SQ SEQUENCE 608 AA; 68914 MW; CF5E92647AAFE9A2 CRC64;

Query Match 90.1%; Score 64; DB 1; Length 608;
Best Local Similarity 85.7%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLHEKTPVSDRVTK 14
Db 486 VLHEKTPVSEKVTK 499

RESULT 9
ALBU RAT
ID ALBU RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN ALB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RC MEDLINE=81233722; PubMed=7017712;
RA Sargent T.D., Yang M., Borner J.;
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
[2]
RP SEQUENCE OF 1-38, AND PROCESSING.
RX MEDLINE=77243657; PubMed=893447;

```

RA Straus A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
RT piece. Analysis of the direct translation product of albumin
RT messenger RNA.";
RL J. Biol. Chem. 252:6846-6855 (1977).
RN [3]
RP SEQUENCE OF 25-222.
RA Isemura S., Ikenaka T.;
RX MEDLINE=78109429; PubMed=564345;
RT "Amino acid sequences of fragments I and II obtained by cyanogen
RT bromide cleavage of rat serum albumin.";
RL J. Biochem. 83:35-48 (1978).
RN [4]
RP SEQUENCE OF 223-288 AND 572-608.
RA Isemura S., Ikenaka T.;
RX MEDLINE=76260153; PubMed=956149;
RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
RT the amino acid sequences of four fragments.";
RL J. Biochem. 79:1183-1196 (1976).
RN [5]
RP SEQUENCE OF 166-174.
RA TISSUE=Plasma;
RX MEDLINE=87194805; PubMed=2437111;
RT Carraway R.E., Mitra S.P., Cochran D.E.;
RT "Structure of a biologically active neurotensin-related peptide
RT obtained from pepsin-treated albumin(s)." ;
RL J. Biol. Chem. 262:5968-5973 (1987).
RN [6]
RP COPPER-BINDING.
RX MEDLINE=79001617; PubMed=80265;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Copper(II)-binding ability of human alpha-fetoprotein." ;
RL Cancer Res. 38:3483-3486 (1978).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and
CC blood flow (potential).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01222; CAA24532.1; -
CC PIR; A93872; ABETS.
CC HSSP; P02768; 1E7B.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608 SERUM ALBUMIN.
FT PEPTIDE 166 174 NEUTROGENIN-RELATED PEPTIDE.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.

FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT VARIANT 282 282 V -> L (IN REF. 5).
FT CONFLICT 174 174 Y -> L (IN REF. 5).
SQ SEQUENCE 608 AA; 68718 MW; 5B497A282411AB7 CRC64;
Query Match 90.1%; Score 64; DB 1; Length 608;
Best Local Similarity 85.1%; Pred. No. 0.0012;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLHEKTPVSDRVTK 14
Db 486 VLHEKTPVSEKVT 499
|||||:|:|:
RESULT 10
ALBU_HORSE STANDARD; PRT; 607 AA.
ID ALBU_HORSE
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)
RT at 0.27-nm resolution.";
RL Eur. J. Biochem. 215:205-212 (1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IGE.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74045; CAA52194.1; -
CC PIR; S34053; ABHOS.
CC HSSP; P02768; 1E7B.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.

```
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 607  
FT DOMAIN 25 204  
FT DOMAIN 211 396  
FT DOMAIN 403 594  
FT METAL 27 27  
FT METAL 77 86  
FT DISULFID 99 115  
FT DISULFID 114 125  
FT DISULFID 147 192  
FT DISULFID 191 200  
FT DISULFID 223 269  
FT DISULFID 268 276  
FT DISULFID 288 302  
FT DISULFID 301 312  
FT DISULFID 339 384  
FT DISULFID 383 392  
FT DISULFID 415 461  
FT DISULFID 460 471  
FT DISULFID 484 500  
FT DISULFID 499 510  
FT DISULFID 537 582  
FT DISULFID 581 590  
SQ SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;  
  
Query Match 88.7%; Score 63; DB 1; Length 607;  
Best Local Similarity 78.6%; Pred. No. 0.0018;  
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VLHEKTPVSDRVTK 14  
|||||:|||||:  
Db 485 VLHEKTPVSEKTK 498  
  
RESULT 11  
ALBU MERUN  
ID _ALBU MERUN STANDARD; PRT; 609 AA.  
AC O35050;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Serum albumin precursor.  
GN ALB.  
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;  
OC Meriones.  
OX NCBI_TaxID=10047;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MGS IDR; TISSUE=Liver;  
RX MEDLINE=9811663; PubMed=9455485;  
RA Yoshida K., Seto-Onshima A., Sinochura H.;  
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic  
synthesis in the Mongolian gerbil, Meriones unguiculatus.";  
RL DNA Res. 4:351-354(1997).  
CC !- FUNCTION: Serum albumin, the main protein of plasma, has a good  
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,  
hormones, bilirubin and drugs. Its main function is the regulation  
of the colloidal osmotic pressure of blood.  
CC !- SUBCELLULAR LOCATION: Secreted.  
CC !- TISSUE SPECIFICITY: Plasma.  
CC !- SIMILARITY: Belongs to the ALB/APP/VDB family.  
CC !- SIMILARITY: Contains 3 albumin domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
DR EMBL; AB006197; BAA21765.1; -.  
DR PIR; JC5838; JC5838.  
DR HSP; P02768; 1E7B.  
DR InterPro; IPR000264; Serum_albumin.  
DR Pfam; PF00273; transport_prot; 3.  
DR PRINTS; PR00802; SERUMALBUMIN.  
DR ProDom; PDO02486; Serum_albumin; 1.  
DR SMART; SM00103; ALBUMIN; 3.  
DR PROSITE; PS00212; ALBUMIN; 3.  
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 609  
FT DOMAIN 25 206  
FT DOMAIN 213 398  
FT DOMAIN 405 596  
FT METAL 28 28  
FT METAL 78 87  
FT DISULFID 100 116  
FT DISULFID 115 126  
FT DISULFID 149 194  
FT DISULFID 193 202  
FT DISULFID 225 271  
FT DISULFID 270 278  
FT DISULFID 290 304  
FT DISULFID 303 314  
FT DISULFID 341 386  
FT DISULFID 385 394  
FT DISULFID 417 463  
FT DISULFID 462 473  
FT DISULFID 486 502  
FT DISULFID 501 512  
FT DISULFID 539 584  
FT DISULFID 583 592  
SQ SEQUENCE 609 AA; 68940 MW; 9CASF97F67EFLA48 CRC64;  
  
Query Match 84.5%; Score 60; DB 1; Length 609;  
Best Local Similarity 78.6%; Pred. No. 0.0062;  
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VLHEKTPVSDRVTK 14  
:|||||:|||||:  
Db 487 LLHEKTPVSEQVTK 500  
  
RESULT 12  
ALBU MOUSE  
ID _ALBU MOUSE STANDARD; PRT; 608 AA.  
AC P07724; Q61802;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Serum albumin precursor.  
GN ALB OR ALB1 OR ALB-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA van Reeth T., Gabant P., Szpirer J., Szpirer C.;  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Tongue;  
RX MEDLINE=21085650; PubMed=11217851;  
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
Arakawa T., Hara A., Fukunishi Y., Komio H., Adachi J., Fukuda S.,
```

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Peole G., Quackenbush J.,
RA Schindl L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Ioyo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Konteuki S.,
RA Hayashizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection.";
RA Nature 409:685-690(2001).
RA [3]
RA SEQUENCE FROM N.A.
RA STRAIN=FVB/N; TISSUE=Liver;
RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Ramer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin I.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length
RA human and mouse cDNA sequences.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RA [4]
RA SEQUENCE OF 99-516 FROM N.A.
RA MEDLINE=86216123; PubMed=2452956;
RA Minghetti P.P., Law S.W., Dugaiczky A.,
RA "The rate of molecular evolution of alpha-fetoprotein approaches that
RA of pseudogenes.";
RA Mol. Biol. Evol. 2:347-358(1985).
RA [5]
RA SEQUENCE OF 477-551 FROM N.A.
RA STRAIN=BALB/C;
RA MEDLINE=90269606; PubMed=1971802;
RA Boccaccio C., Deschatrete J., Meunier-Rotival M.,
RA "Empty and occupied insertion site of the truncated LINE-1 repeat
RA located in the mouse serum albumin-encoding gene.";
RA Gene 88:181-186(1990).
RA [6]
RA SEQUENCE OF 25-44.
RA TISSUE=Liver;
RA MEDLINE=93162044; PubMed=1286668;
RA Giometti C.S., Taylor J., Tollaksen S.L.,
RA "Mouse liver protein database: a catalog of proteins detected by two-
RA dimensional gel electrophoresis.";
RA Electrophoresis 13:970-991(1992).
RA -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
RA binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
RA hormones, bilirubin and drugs. Its main function is the regulation
RA of the colloidal osmotic pressure of blood.
RA -!- SUBCELLULAR LOCATION: Secreted.
RA -!- TISSUE SPECIFICITY: Plasma.
RA -!- SIMILARITY: Belongs to the ALB/AFP/VDS family.
RA -!- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ011413; CAA09617.1; -;
CC EMBL; AK010025; BAB26650.1; -;
CC EMBL; BC049971; AAA49971.1; -;
CC EMBL; M16111; AAA37190.1; -;
CC EMBL; X13060; CAA31458.1; -;
CC PIR; A05139; A05139.
CC HSSP; P02768; 1E78.
CC SWISS-2DPAGE; P0724; MOUSE.
CC MGD; MGI:87991; Alb1.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PRO0802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN_3.
CC PROSITE; PS00212; ALBUMIN; 3. Repeat; Signal; Copper.
CC Metal-binding; Lipid-binding; BY SIMILARITY.
CC SIGNAL 1 18
CC FT PROPEP 19 24 SERUM ALBUMIN.
CC FT CHAIN 25 608 ALBUMIN 1.
CC FT DOMAIN 25 205 ALBUMIN 2.
CC FT DOMAIN 212 397 ALBUMIN 3.
CC FT DOMAIN 404 595 COPPER.
CC FT METAL 27 27 BY SIMILARITY.
CC FT DISULFID 77 86 BY SIMILARITY.
CC FT DISULFID 99 115 BY SIMILARITY.
CC FT DISULFID 114 125 BY SIMILARITY.
CC FT DISULFID 148 193 BY SIMILARITY.
CC FT DISULFID 192 201 BY SIMILARITY.
CC FT DISULFID 224 270 BY SIMILARITY.
CC FT DISULFID 269 277 BY SIMILARITY.
CC FT DISULFID 289 303 BY SIMILARITY.
CC FT DISULFID 302 313 BY SIMILARITY.
CC FT DISULFID 340 385 BY SIMILARITY.
CC FT DISULFID 384 393 BY SIMILARITY.
CC FT DISULFID 416 462 BY SIMILARITY.
CC FT DISULFID 461 472 BY SIMILARITY.
CC FT DISULFID 485 501 BY SIMILARITY.
CC FT DISULFID 500 511 BY SIMILARITY.
CC FT DISULFID 538 583 BY SIMILARITY.
CC FT DISULFID 582 591 BY SIMILARITY.
CC FT CONFLICT 27 27 H -> D (IN REF. 6).
CC FT CONFLICT 33 33 H -> D (IN REF. 6).
CC FT CONFLICT 41 41 Q -> I (IN REF. 6).
CC SQ SEQUENCE 608 AA; 68692 MW; 292F7C7ED3A61B4 CRC64;
Query Match 83.1%; Score 59; DB 1; Length 608;
Best Local Similarity 78.6%; Pred. No. 0.0093;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 VLHEKTPVSDRVTK 14
Db 486 LLHEKTPVSEHVK 499
RESULT 13
FLHA_HELPJ STANDARD; PRT; 733 AA.
ID FLHA_HELPJ
AC Q9ZM40;
DT 16-OCT-2001 (Rel. 40, Created)
DF 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar biosynthesis protein flha.
DE FLHA OR FLBA OR JHP0383.
GN Helicobacter pylori J99 (Campylobacter pylori J99).
OS

```

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OC NCBI_TaxID=85963;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RX ALM R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Carubeo A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RA "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RT Nature 397:176-180(1999).
CC -1- FUNCTION: INVOLVED IN THE EXPORT OF FLAGELLUM PROTEINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE FHIPEP (FLAGELLA/HR/INVASION PROTEINS
CC EXPORT PORE) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sb-sib.ch).
-----
CC EMBL; AE001473; AAD05964.1; -
CC PIR; E71937; E71937.
CC InterPro; IPR001712; Bact_FHIPEP.
CC InterPro; IPR006301; FlHa.
CC Pfam; PF00771; FHIPEP; 1.
CC PRINTS; PR00949; TYPE3IMAPROT.
CC TIGRFS; TIGR01398; FlHa; 1.
CC PROSITE; PS00994; FHIPEP; FALSE NEG.
KM Flagellum; Transport; Protein transport; Transmembrane;
KW Complete proteome.
FT FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 49 69 POTENTIAL.
FT TRANSMEM 125 145 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 291 311 POTENTIAL.
FT TRANSMEM 313 333 POTENTIAL.
SQ SEQUENCE 733 AA; 80778 MW; EDB00DF3A2F28F0E CRC64;
Query Match 60.6%; Score 43; DB 1; Length 733;
Best Local Similarity 53.8%; Pred. No. 7.7;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VLHEKTPVSDRVT 13
Db :|||:|:|
577 LLHEKIPKIDMLT 589

RESULT 14
FLHA_HELPY STANDARD; PRT; 733 AA.
AC O06758; O06759;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar biosynthesis protein flha.
GN FLHA OR FLHA OR HPI041.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=859;
RA Schmitz A., Josenhans C., Suerbaum S.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

```



```

Query Match      59.2%; Score 42; DB 1; Length 905;
Best Local Similarity 77.8%; Pred. NO. 14;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 HEKTPVSDR 11
      |:|||||
Ddb      592 HOKTFTSDR 600

Search completed: April 19, 2004, 11:52:56
Job time : 2.02124 secs

```

GenCore version 5.1.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 5.40351 Seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_462_475

Perfect score: 71

Sequence: 1 VLHEKTPVSDRVTK 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriaph.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	71	100.0	396	4	Q81UK7
2	71	100.0	417	4	Q86YGO
3	67	94.4	584	6	Q7YSG3
4	63	88.7	608	5	Q95VB7
5	59	83.1	576	11	Q8C7C7
6	59	83.1	608	11	Q8C7H3
7	46	64.8	148	10	Q8GSH6
8	43	60.6	553	16	Q91783
9	43	60.6	733	2	Q86G93
10	42	59.2	905	16	Q87ZQ4
11	41	57.7	95	2	Q93EH4
12	41	57.7	231	16	Q8F9E1
13	41	57.7	735	16	Q7VIY5
14	40	56.3	358	2	Q8VUT7
15	40	56.3	604	17	Q97YJ0
16	40	56.3	630	10	Q9SF07

17	40	56.3	652	4	Q8N990
18	40	56.3	699	5	O01545
19	40	56.3	3456	12	P89201
20	39	54.9	63	16	Q989J5
21	39	54.9	137	2	Q8KZS5
22	39	54.9	385	16	Q92KLO
23	39	54.9	577	2	O33047
24	39	54.9	682	17	Q9HM85
25	39	54.9	753	16	Q9CBE1
26	39	54.9	1125	5	Q9NFA7
27	38	53.5	172	16	O69637
28	38	53.5	172	16	O7TVZ5
29	38	53.5	309	2	P72172
30	38	53.5	309	5	Q9NUL1
31	38	53.5	309	16	Q9HUZ7
32	38	53.5	333	16	Q8DTB0
33	38	53.5	375	16	Q8A079
34	38	53.5	444	10	Q8S9T6
35	38	53.5	457	10	O82108
36	38	53.5	577	2	Q9ALJ1
37	38	53.5	690	16	Q97H66
38	38	53.5	768	2	Q84FP3
39	38	53.5	777	2	Q8GM52
40	38	53.5	931	16	Q82K00
41	38	53.5	1522	5	Q7YSM9
42	38	53.5	1771	3	O870V1
43	37	52.1	131	6	Q8HXG0
44	37	52.1	161	16	Q8F0A5
45	37	52.1	216	3	Q9USA7

ALIGNMENTS

RESULT 1
Q81UK7 PRELIMINARY; PRT; 396 AA.
AC Q81UK7;
CT 01-11-2003 (TRENBLrel. 23, Created)
DT 01-11-2003 (TRENBLrel. 23, Last sequence update)
DT 01-11-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] SEQUENCE FROM N.A.
RP TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC035969; AAH35969.1; .
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0008610; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 2.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 71; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 273 VLHEKTPVSDRVTK 286

RESULT 2
Q86YGO

```

ID Q86YGO PRELIMINARY; PRT; 417 AA.
AC Q86YGO;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47350 MW; 16E764833EEF4E8D CRC64;

Query Match 100.0%; Score 71; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 294 VLHEKTPVSDRVTK 307
|||||
|||||

RESULT 3
Q7YSG3 PRELIMINARY; PRT; 584 AA.
AC Q7YSG3;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Albumin (Fragment).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R.; Swoboda I.; Bohle B.; Hauswirth A.W.; Valent P.;
RA Rumpold H.; Valenta R.; Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
RT albumin: IGE recognition, induction of basophil activation and
RT lymphoproliferative responses in atopic patients.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487677; CAD32275.1; -.
FT NON_TER 1
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 94.4%; Score 67; DB 6; Length 584;
Best Local Similarity 92.9%; Pred. No. 0.0016;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 462 VLHEKTPVSDRVTK 475
|||||
|||||

RESULT 4
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A.; Asahi H.; Stadecker M.J.; LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RT component in murine infection with Schistosoma mansoni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418550; AAL08579.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EAB28E1C6E54 CRC64;

Query Match 88.7%; Score 63; DB 5; Length 608;
Best Local Similarity 85.7%; Pred. No. 0.0086;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
DB 486 VLHEKTPVSDRVTK 499
|||||
|||||

RESULT 5
Q8C7C7 PRELIMINARY; PRT; 576 AA.
AC Q8C7C7;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050844; BAC34360.1; -.
DR MGD; MGI:187991; Albl.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON_TER 1
SQ SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;

```

```

Query Match      83.1%; Score 59; DB 11; Length 576;
Best Local Similarity 78.6%; Pred. No. 0.043; 1; Indels 0; Gaps 0;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 454 LLHEKTPVSEHVTK 467

RESULT 6
Q8C7H3 PRELIMINARY; PRT; 608 AA.
AC Q8C7H3;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Albumin 1.
GN ALBL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI:87991; Albl.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;

Query Match      83.1%; Score 59; DB 11; Length 608;
Best Local Similarity 78.6%; Pred. No. 0.046; 1; Indels 0; Gaps 0;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 486 LLHEKTPVSEHVTK 499

RESULT 7
Q8GSH6 PRELIMINARY; PRT; 148 AA.
AC Q8GSH6;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE QJ1477_F01.19 protein (P0496C02.4 protein).
GN QJ1477_F01.19 OR P0496C02.4.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:OJ1477_F01.",

```

```

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
RT clone:P0496C02.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003833; BAC15482.1; -.
DR EMBL; AP004378; BAC16046.1; -.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 2.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS00088; ANK_REPEAT; 2.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
SQ SEQUENCE 148 AA; 15284 MW; 1D69B65DE09D4364 CRC64;

Query Match      64.8%; Score 46; DB 10; Length 148;
Best Local Similarity 66.7%; Pred. No. 2.5;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 HEKTPVSDRVTK 14
Db 113 HEKTPVSDRVTK 124

RESULT 8
Q9I783 PRELIMINARY; PRT; 553 AA.
ID Q9I783;
AC Q9I783;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein PA0049.
GN PA0049.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Muzoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Geltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004444; AAG03439.1; -.
DR PIR; D83640; D83640.
DR GO; GO:0004219; F:pyroglutamy-peptidase I activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000816; Peptidase_C15.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 553 AA; 60113 MW; BE9B7E723D46FB3B CRC64;

Query Match      60.6%; Score 43; DB 16; Length 553;
Best Local Similarity 42.9%; Pred. No. 33;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVTK 14
Db 71 VTHDQTPLEAERIVK 84

RESULT 9
Q9S6G9 PRELIMINARY; PRT; 733 AA.
ID Q9S6G9
AC Q9S6G9;

```

DR 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Flagellar biosynthesis/regulation protein FliA.
GN FLA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43504;
RX MEDLINE=99214098; PubMed=10198012;
RA McGee D.J., May C.A., Garner R.M., Himpel J.M., Mobley H.L.T.;
RT "Isolation of Helicobacter pylori genes that modulate urease
activity";
RL J. Bacteriol. 181:2477-2484 (1999).
DR EMBL; AF123197; AAD27689.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR InterPro; IPR001712; Bact_FHIPSP.
DR InterPro; IPR006301; FLA-
DR Pfam; PF00771; FHIPEP; 1.
DR PRINTS; PRO0949; TYP3IAPROT.
DR TIGRFAMs; TIGR01398; FLA; 1.
SQ SEQUENCE 733 AA; 80883 MW; 6E8817AB4268128B CRC64;

Query Match 60.6%; Score 43; DB 2; Length 733;
Best Local Similarity 53.8%; Pred. No. 43;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVT 13
:|||||:|:
DB 577 LLHEKIPKDMLT 599

RESULT 10
Q87ZQ4
ID Q87ZQ4 PRELIMINARY; PRT; 905 AA.
AC Q87ZQ4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase I, G subunit.
GN NUOG OR PSPT03370.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Uterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,
RA White O., Fraser C., Collier A.;
RT "Complete sequence of Pseudomonas syringae";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016968; AA056948.1; -;
DR TIGR; PST03370; -;
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR009010; Asp decarb fold.
DR InterPro; IPR000283; Complex 1 75K.
DR InterPro; IPR001041; Ferridoxin.
DR InterPro; IPR006656; Molybdopterin.
DR InterPro; IPR006963; Molybdop_Fe4S4.
DR Pfam; PF00111; fer2; 1.
DR Pfam; PF00384; molybdopterin; 1.
DR Pfam; PF04879; Molybdop_Fe4S4; 1.

DR PROSITE; PS00641; COMPLEX1_75K_1; 1.
DR PROSITE; PS00642; COMPLEX1_75K_2; 1.
DR PROSITE; PS00643; COMPLEX1_75K_3; 1.
KW Complete proteome.
SQ SEQUENCE 905 AA; 98333 MW; 53716244354E40DD CRC64;

Query Match 59.2%; Score 42; DB 16; Length 905;
Best Local Similarity 77.8%; Pred. No. 81;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 HEKTPVSDR 11
:|||||:
DB 592 HQTPTSDR 600

RESULT 11
Q93EH4
ID Q93EH4 PRELIMINARY; PRT; 95 AA.
AC Q93EH4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Flagellar biosynthesis protein (fragment).
GN FLHA.
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3B1;
RA Ge Z., Feng Y., Fox J.G.;
RT "Helicobacter hepaticus genome: construction of an ordered cosmid
library and sequence analysis of the selected genomic regions";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF358691; AAL16680.1; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0009306; P:protein secretion; IEA.
DR InterPro; IPR001712; Bact_FHIPEP.
DR Pfam; PF00771; FHIPEP; 1.
DR NON_TER 1
FT NON_TER 95
FT NON_TER 95
SQ SEQUENCE 95 AA; 10789 MW; BD44A1BL53EBA53 CRC64;

Query Match 57.7%; Score 41; DB 2; Length 95;
Best Local Similarity 53.8%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 VLHEKTPVSDRVT 13
:|||||:|:
DB 51 LLHEKVPKDMPT 63

RESULT 12
Q8F9B1
ID Q8F9B1 PRELIMINARY; PRT; 231 AA.
AC Q8F9B1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN LA0284.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011216; AAN47483.1; -;
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 231 AA; 26471 MW; 15024FDS6DE290D9 CRC64;

[illegible]

ng

IT Italy";
Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
[22]
VARIANT VENEZIA.
MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
RT carboxyl-terminal variants of human serum albumin.";
Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
[23]
VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RA KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Mateuda Y.-I., Anaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese.";
Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
[24]
VARIANT CASEBOOK ASN-518.
MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
RT albumin: albumin Casebrook (494 Asp-->Asn).";
Biochim. Biophys. Acta 1097:49-54(1991).
[25]
VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two allalbumins with identical electrophoretic mobility are produced
Query Match 100.0%; Score 33; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HGDLLLE 6
Db 271 HGDLLLE 276
RESULT 15
ALBU MERUN
ID ALBU MERUN STANDARD; PRT; 609 AA.
AC Q35090;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OX NCBI_TaxID=10047;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MGS IDR; TISSUE=Liver;
RX MEDLINE=98116663; PubMed=9455485;
RA Yoshida K., Seto-Oshima A., Sinohara H.;
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
RT synthesis in the Mongolian gerbil, Meriones unguiculatus.";
RNA Res. 4:351-354(1997).
RL -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)

CC EMBL; AB006197; BAA21765.1; -;
DR PIR; JCS838; JCS838.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport p1ct; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 609 SERUM ALBUMIN.
FT DOMAIN 25 206 ALBUMIN 1.
FT DOMAIN 213 398 ALBUMIN 2.
FT DOMAIN 405 596 ALBUMIN 3.
FT METAL 28 28 COPPER.
FT DISULFID 78 87 BY SIMILARITY.
FT DISULFID 100 116 BY SIMILARITY.
FT DISULFID 115 126 BY SIMILARITY.
FT DISULFID 143 194 BY SIMILARITY.
FT DISULFID 193 202 BY SIMILARITY.
FT DISULFID 225 271 BY SIMILARITY.
FT DISULFID 270 278 BY SIMILARITY.
FT DISULFID 290 304 BY SIMILARITY.
FT DISULFID 303 314 BY SIMILARITY.
FT DISULFID 341 386 BY SIMILARITY.
FT DISULFID 385 394 BY SIMILARITY.
FT DISULFID 417 463 BY SIMILARITY.
FT DISULFID 462 473 BY SIMILARITY.
FT DISULFID 486 502 BY SIMILARITY.
FT DISULFID 501 512 BY SIMILARITY.
FT DISULFID 539 584 BY SIMILARITY.
FT DISULFID 583 592 BY SIMILARITY.
SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F67EF.A48 CRC64;
Query Match 100.0%; Score 33; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HGDLLLE 6
Db 272 HGDLLLE 277
Search completed: April 19, 2004, 11:52:50
Job time : 0.437673 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 2.31579 Seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_247_252

Perfect score: 33
Sequence: 1 HGDLLLE 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	123	11 Q7TQM6	Q7TQM6 meriones un
2	33	100.0	129	6 Q8WMJ3	Q8wmj3 macaca mulla
3	33	100.0	194	6 Q8WMB8	Q8wmb8 macaca mulla
4	33	100.0	211	6 O19044	O19044 oryctolagus
5	33	100.0	250	16 Q91612	Q91612 pseudomonas
6	33	100.0	251	16 Q84712	Q84712 pseudomonas
7	33	100.0	283	2 Q8RBV8	Q8rbv8 pseudomonas
8	33	100.0	379	4 Q8N417	Q8n417 homo sapien
9	33	100.0	379	11 Q8C725	Q8c725 mus musculus
10	33	100.0	417	4 Q85YGO	Q85ygo homo sapien
11	33	100.0	419	16 Q7U5E3	Q7u5e3 synchococc
12	33	100.0	423	13 Q98975	Q98975 oncorhynch
13	33	100.0	529	16 Q9KYCA	Q9kyc4 streptomyce
14	33	100.0	560	2 Q8KQ21	Q8kq21 pseudomonas
15	33	100.0	560	16 Q8BD23	Q8bd23 pseudomonas
16	33	100.0	576	11 Q8C7C7	Q8c7c7 mus musculus

Q7yeg3 felis silve
Q95vb7 schistosoma
Q8c7h3 mus musculus
Q8fw79 bruceella su
Q8yc46 bruceella me
Q9b968 ceratocolen
Q9yfr4 aeropyrum p
Q89ck9 bradyrhizob
Q8j44 xenopus lae
Q892r2 clostridium
Q9tni6 sphagnum fa
Q8c0i9 mus musculus
Q8c1l9 mus musculus
Q9rv18 deinococcus
Q86w83 homo sapien
O16677 caenorhabdi
Q9r9w4 helicobacte
Q9r9w4 pseudomonas
Q923y1 pseudomonas
Q8rv58 pseudomonas
Q8kq23 pseudomonas
Q93mw5 burkholderi
Q7x5k2 pseudomonas
Q88d25 pseudomonas
Q8nck8 homo sapien
Q8n2h2 homo sapien
Q8bz03 mus musculus
Q8kiy2 mus musculus
Q93pd2 azoarcus sp

ALIGNMENTS

RESULT 1

Q7TQM6 PRELIMINARY; PRT; 123 AA.

AC Q7TQM6;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Inward-rectifying potassium channel KCNJ10 (Fragment).
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OX NCBJ_TaxID=10047;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Wangemann P., White E.M., Albrecht B., Wu T., Maganti R.J.,
RA Jabba S.V., Lee J.H., Everett L.A., Royaux I.E., Green E.D.,
RA Marcus D.C.;
RT "Loss of KCNJ10 expression abolishes endocochlear potential and causes
RT deafness in Pendreds syndrome mouse model."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY308017; AAF74570.1; -
KW Ionic channel.

FT NON_TER 123 123
FT NON_TER 123 123
SQ SEQUENCE 123 AA; 13412 MW; EDD82F98F8908A4E CRC64;

Query Match 100.0%; Score 33; DB 11; Length 123;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6

DB 4 HGDLLLE 9

RESULT 2

Q8WMJ3

RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Raizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an
 RT opportunistic pathogen."
 RL Nature 406:959-964 (2000).
 DR ENBL; AE004466; AAG03667.1; -.
 DR PIR; F83609; F83609. DUF81.
 DR InterPro; IPR002781; DUF81.
 DR Pfam; PF01925; DUF81; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 250 AA; 25773 MW; C5F5D079CEDA49A1 CRC64;

Query Match 100.0%; Score 33; DB 16; Length 250;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
 |||||
 Db 191 HGDLL 196

RESULT 6
 Q84712 PRELIMINARY; PRT; 251 AA.
 AC Q84712
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Pseudomonas putida.
 OG Plasmid pDTG1.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIB 9816-4;
 RA Zylstra G.J., Dennis J.J.;
 RT "Complete nucleotide sequence of the NAH plasmid pDTG1 from
 RT Pseudomonas putida NCIB 9816-4."
 RL Submitted (DSC-2002) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AF491307; AA064230.1; -.
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 DR InterPro; IPR006015; USP.
 DR PRINTS; PR01438; UNYRS1STRESS.
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 251 AA; 27295 MW; F9F4CEEFFB187FA5 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 251;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
 |||||
 Db 73 HGDLL 78

RESULT 7
 Q9REVB PRELIMINARY; PRT; 283 AA.
 AC Q9REVB
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Pseudomonas sp. R9.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=101164;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R9; TRANSPOSON=Tn1404;

RX MEDLINE=20011227; PubMed=10543801;
 RA Schnabel E.L., Jones A.L.;
 RT "Distribution of tetracycline resistance genes and transposons among
 RT phytoplane bacteria in Michigan apple orchards."
 RL Appl. Environ. Microbiol. 65:4898-4907 (1999).
 DR ENBL; AF157797; AAD47991.1; -.
 DR GO; GO:0006950; P:response to stress; IEA.
 DR InterPro; IPR006016; Usp_dom.
 DR Pfam; PF00582; Usp; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 283 AA; 30670 MW; 5586DD9392524DDE CRC64;

Query Match 100.0%; Score 33; DB 2; Length 283;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
 |||||
 Db 105 HGDLL 110

RESULT 8
 Q8N417 PRELIMINARY; PRT; 379 AA.
 AC Q8N417
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Potassium inwardly-rectifying channel, subfamily J, member 10.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR ENBL; SC034036; AAH34035.1; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005242; P:forward rectifier potassium channel activity; IEA.
 DR GO; GO:0005267; P:potassium channel activity; IEA.
 DR GO; GO:0006813; P:potassium ion transport; IEA.
 DR InterPro; IPR001838; K-channel_IR.
 DR InterPro; IPR001522; K-channel_pore.
 DR Pfam; PF01007; IRK; 1.
 DR PRINTS; PR01320; KIRCHANNEL.
 DR ProDom; PD001103; K-channel_IR; 2.
 SQ SEQUENCE 379 AA; 42550 MW; 9D8E156BECEB4CDD CRC64;

Query Match 100.0%; Score 33; DB 4; Length 379;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
 |||||
 Db 93 HGDLL 98

RESULT 9
 Q8C7Z5 PRELIMINARY; PRT; 379 AA.
 AC Q8C7Z5
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Potassium inwardly-rectifying channel.
 GN KCNJ10.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK048864; BAC33477.1; -.
DR MGD; NGI:1194504; Kcnj10.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005242; F:inward rectifier potassium channel activity; IEA.
DR GO; GO:0005267; F:potassium channel activity; IEA.
DR GO; GO:0006813; P:potassium ion transport; IEA.
DR InterPro; IPR001838; K-channel_IR.
DR InterPro; IPR001622; K-channel_pore.
DR Pfam; PF01007; IRK; 1.
DR PRINTS; PR01320; KIRCHANNEL.
DR ProDom; PD001103; K-channel_IR; 2.
SQ SEQUENCE 379 AA; 42459 MW; 7FF6E1360F62E9B3 CRC64;

Query Match 100.0%; Score 33; DB 11; Length 379;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
Db 93 HGDLL 98

RESULT 10
Q86YGO PRELIMINARY; PRT; 417 AA.
AC Q86YGO;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_Prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47360 MW; 16E764833EEF4E8D CRC64;

Query Match 100.0%; Score 33; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
Db 79 HGDLL 84

RESULT 11
Q7U5E3 PRELIMINARY; PRT; 419 AA.
ID Q7U5E3;
AC Q7U5E3;
DT 01-OCT-2003 (TRENBLrel. 25, Created)

DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative dihydroorotase (EC 3.5.2.3).
GN SYNW1764.
OS Synchococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641;
RA Palenik B., Brahamsa B., Larimer F.W., Land M., Hauser L., Chain P.,
RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,
RA Lufresne A., Patensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synchococcus";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569693; CAE08279.1; -.
DR Hydrolase; Complete proteome.
KW Hydrolase;
SQ SEQUENCE 419 AA; 44483 MW; DFF0860393F3298B CRC64;

Query Match 100.0%; Score 33; DB 16; Length 419;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
Db 142 HGDLL 147

RESULT 12
Q98975 PRELIMINARY; PRT; 423 AA.
AC Q98975;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Weakly inward rectifying potassium channel.
GN SWIRK.
OS Oncorhynchus masou (Cherry salmon) (Masu salmon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8020;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96279099; PubMed=8663136;
RA Kubo Y., Miyashita T., Kubokawa K.;
RT "A weakly inward rectifying potassium channel of the salmon brain:
RT Glutamate 179 in the second transmembrane domain is insufficient for
RT strong rectification";
RL J. Biol. Chem. 271:15729-15735(1996).
DR EMBL; D83537; BAA12009.2; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005242; F:inward rectifier potassium channel activity; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0005267; P:potassium channel activity; IEA.
DR GO; GO:0006813; P:potassium ion transport; IEA.
DR InterPro; IPR001838; K-channel_IR.
DR InterPro; IPR001622; K-channel_pore.
DR Pfam; PF01007; IRK; 1.
DR PRINTS; PR01320; KIRCHANNEL.
DR ProDom; PD001103; K-channel_IR; 2.
KW Ionic channel.
SQ SEQUENCE 423 AA; 46950 MW; CDDF0159275EAE79 CRC64;

Query Match 100.0%; Score 33; DB 13; Length 423;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
Db 111 HGDLL 116

RESULT 13

Q9KYC4 PRELIMINARY; PRT; 529 AA.
 AC Q9KYC4;
 DT 01-OCT-2000 (TREMELrel. 15, Created)
 DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE Hypothetical protein SCO6906.
 GN SCO6906 OR SC1B2.12C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdono-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Lark L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL939129; CAB92568.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 529 AA; 57433 MW; 87075B0602262E11 CRC64;

Query Match 100.0%; Score 33; DB 16; Length 529;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6
 |||||
 Db 182 HGDLLE 187

RESULT 14

Q8KQ21 PRELIMINARY; PRT; 560 AA.
 AC Q8KQ21;
 DT 01-OCT-2002 (TREMELrel. 22, Created)
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE PHA synthase 2.
 GN PHA2.
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Lee Y.-H., Kim T.-K., Shin H.-D., Lee J.-N., Seo M.-C.;
 RT "Molecular structure of PCR cloned PHA synthase I gene of Pseudomonas
 putida KT2440 and its utilization for medium chain length
 polyhydroxyalkanoate production.";
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY113181; ZAM63409.1; -
 DR GO; GO:0003993; Fracid phosphatase activity; IEA.
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000560; HisAc_pheptase.
 DR Pfam; PF00561; abhydrolase; 1.
 DR PROSITE; PS00616; HIS ACID_PHOSPHAT_1; 1.
 SQ SEQUENCE 560 AA; 62640 MW; COA768C17569B4D3 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 560;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6
 |||||
 Db 410 HGDLLE 415

RESULT 15

Q88D23 PRELIMINARY; PRT; 560 AA.
 AC Q88D23;
 DT 01-JUN-2003 (TREMELrel. 24, Created)
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Poly(3-hydroxyalkanoate) polymerase 2.
 GN PHAC OR PP5005.
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
 RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
 RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
 RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
 RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
 RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuermmier B.,
 RA Frazer C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808(2002).
 DR EMBL; AF016792; AAN70571.1; -
 DR TIGR; PF05005; -; Fracid phosphatase activity; IEA.
 DR GO; GO:0003993; Fracid phosphatase activity; IEA.
 DR InterPro; IPR000073; A/b hydrolase.
 DR InterPro; IPR000560; HisAc_pheptase.
 DR Pfam; PF00561; abhydrolase; 1.
 DR PROSITE; PS00616; HIS ACID_PHOSPHAT_1; 1.
 KW Complete proteome.
 SQ SEQUENCE 560 AA; 62722 MW; 4BF0D75DB5874868 CRC64;

Query Match 100.0%; Score 33; DB 16; Length 560;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6
 |||||
 Db 410 HGDLLE 415

Search completed: April 19, 2004, 12:00:07
 Job time : 4.31579 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 3.52909 Seconds
(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_247_252

Perfect score: 33
Sequence: 1 HGDLL6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	14	ABB52153	Human API
2	33	100.0	14	ABP61551	Human KRP
3	33	100.0	14	ABR58983	Alzheimer
4	33	100.0	14	ABR75499	Liver res
5	33	100.0	18	ABR12469	Human alb
6	33	100.0	21	ABR62326	Bovine se
7	33	100.0	58	ABB40958	Peptide #
8	33	100.0	58	AAW34732	Peptide #
9	33	100.0	58	ABB25068	Protein #
10	33	100.0	58	AAW74619	Human bon
11	33	100.0	58	AAW61818	Human bra
12	33	100.0	58	ABG56405	Human liv
13	33	100.0	58	ABG44448	Human pep
14	33	100.0	85	AAW18788	Peptide #
15	33	100.0	85	ABB37889	Peptide #
16	33	100.0	85	AAW31296	Peptide #
17	33	100.0	85	ABB23147	Protein #
18	33	100.0	85	AAW71019	Human bon
19	33	100.0	85	AAW58521	Human bra
20	33	100.0	85	ABG52734	Human liv
21	33	100.0	85	ABG40810	Human pep
22	33	100.0	114	AAW25284	Human pro
23	33	100.0	116	AAW00108	Human pol
24	33	100.0	123	AAW23861	abCEST en
25	33	100.0	126	AAW04435	Human pol

26	33	100.0	131	4	ABG01238	Novel hum
27	33	100.0	143	4	AAO02571	Human pol
28	33	100.0	188	3	AAW83948	Yeast cod
29	33	100.0	265	7	AAW25672	Dog aller
30	33	100.0	265	7	ADC34920	Dog aller
31	33	100.0	293	4	AAU33082	Novel hum
32	33	100.0	303	2	AAW14178	Human ser
33	33	100.0	365	5	ABP51353	Human MDD
34	33	100.0	373	1	AAW90387	N-termina
35	33	100.0	379	2	AAW26368	Human kid
36	33	100.0	379	5	AAW47972	Human KIR
37	33	100.0	379	5	AAW47973	Human KIR
38	33	100.0	379	7	ADE31731	Human 946
39	33	100.0	388	1	AAW90389	N-termina
40	33	100.0	389	1	AAW90390	N-termina
41	33	100.0	390	1	AAW90391	N-termina
42	33	100.0	407	1	AAW90392	N-termina
43	33	100.0	463	2	AAW14179	Human ser
44	33	100.0	500	7	ADD32019	Heterolog
45	33	100.0	550	4	AAU29877	Novel hum

ALIGNMENTS

RESULT 1

ABB52153
ID ABB52153 standard; peptide; 14 AA.

XX ABB52153;
DT 08-FEB-2002 (first entry)
DE Human API-95 tryptic digest peptide #1.

XX Human; neuroprotective; nootropic; gene therapy; vaccine;
KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;
KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;
KW Expression Reference Protein Isoform; ERPI; proteolysis.

OS Homo sapiens.

XX WO200175454-A2.

XX 11-OCT-2001.

XX 03-APR-2001; 2001WO-US010908.

XX 03-APR-2000; 2000US-0194504P.

XX 28-NOV-2000; 2000US-0253647P.

XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX (PFIZ) PFIZER INC.

XX Durham KL, Friedman DL, Herath HM, Kimmel LH, Parekh RB;

XX Potter DM, Rohlf C, Silber BM, Stiger TR, Sunderland PT;

XX Townsend RR, White F, Williams SA;

XX WPI; 2001-639384/73.

XX Screening for Alzheimer's disease in a mammal, by making two-dimensional

PT array of a feature whose relative abundance correlates with disease, and
PT comparing with abundance of the feature in samples of healthy persons.
XX Example; Page 29; 162pp; English.

XX The invention relates to methods for the screening, diagnosis and

CC prognosis of Alzheimer's disease. The methods involve the detection of

CC Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-

CC Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or

CC plasma. The abundance of the AFs and APIs is then normalised to an

CC Expression Reference Protein Isoform (ERPI) in order to determine whether

CC a patient is suffering from, or has a predisposition to, Alzheimer's

CC Disease. The relative abundance of the Afs and APIs correlates with the
 CC severity of Alzheimer's Disease. The present sequence is a peptide
 CC produced from an API by proteolysis
 XX
 SQ Sequence 14 AA;

Query Match 100.0%; Score 33; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGDLLLE 6
 DB 4 HGDLLLE 9

RESULT 2

ABP61551
 ID ABP61551 standard; peptide; 14 AA.

XX AC ABP61551;

DT 02-OCT-2002 (first entry)

DE Human KRPI tryptic digest peptide #46.

XX Human; tryptic digest peptide; KRPI; kidney response; KR;
 KW nephrotropic kidney response-associated protein isoform; gene therapy;
 KW antitense therapy; kidney function; tubular nephritis; renal failure;
 KW nephron cell metabolic pathway modulation; glomerular necrosis;
 KW papillary necrosis.

XX OS Homo sapiens.

XX PN WO200254081-A2.

XX PD 11-JUL-2002.

XX PF 24-DEC-2001; 2001WO-GB005777.

XX PR 29-DEC-2000; 2000US-0260392P.

XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX PI Holt GD, Kelly MD, Kennedy SJ, Moyses C;

XX DR WPI; 2002-583637/62.

XX Screening, diagnosis or prognosis of kidney response in subject, by
 PT detecting kidney response-associated features or kidney response-
 PT associated protein isoforms in body fluid or tissue from subject.

XX PS Disclosure; Page 40; 168pp; English.

XX The invention relates to a novel method for the screening, diagnosis or
 CC prognosis of kidney response (KR). The method of the invention has
 CC nephrotropic activity, and may have a use in gene therapy or antisense
 CC therapy. The method is useful for the screening, diagnosis or prognosis
 CC of KR in a subject, for determining the stage or severity of KR in a
 CC subject, for identifying a subject at risk of developing KR, or for
 CC monitoring the effect of therapy administered to a subject with KR. An
 CC alternative method of the invention is useful for screening agents that
 CC interact with one or more of the kidney response-associated protein
 CC isoforms (KRPIs). The kidney response includes alterations in kidney
 CC function, any phase of nephron cell metabolic pathway modulation,
 CC glomerular/proximal tubular nephritis, glomerular/papillary necrosis,
 CC acute and chronic renal failure, and end stage renal disease. The
 CC sequences shown in ABP61514-ABP61787 represent tryptic digest peptides of
 CC the KRPIs of the invention.

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 33; DB 5; Length 14;
 Best Local Similarity 100.0%; Pred. No. 5.5;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HGDLLLE 6
 DB 4 HGDLLLE 9

RESULT 3

ABR58983
 ID ABR58983 standard; peptide; 14 AA.

XX AC ABR58983;

DT 11-JUL-2003 (first entry)

XX Alzheimer's Disease-associated protein isoform, AFI-95, SEQ ID 77.

XX Nootropic; Neuroprotective; Alzheimer's disease; API; human;

XX Alzheimer's Disease-associated protein isoform.

XX OS Homo sapiens.

XX PN WO2003028543-A2.

XX PD 10-APR-2003.

XX PF 03-OCT-2002; 2002WO-US031642.

XX PR 03-OCT-2001; 2001US-0326708P.

XX PA (PFIZ) PFIZER PROD INC.

XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX Durham IK, Friedman DL, Herath HM, Kimmel LH, Parekh RB;

XX Potter DM, Rohlf C, Silber SM, Snyder PJ, Soares HD, Stieger TR;

XX Sunderland PT, Townsend RR, White WF, Williams SA;

XX WPI; 2003-371957/35.

XX Screening or diagnosing of Alzheimer's disease (AD) determine the stage
 PT or severity of AD in a subject, comprises analyzing a test sample of body
 PT fluid from the subject by 2-dimensional electrophoresis.

XX Claim 2; Page 43; 179pp; English.

XX The present invention relates to methods for screening or diagnosing
 CC Alzheimer's disease (AD) to determine the stage or severity of AD in a
 CC subject to identify subject at risk of developing AD, or to monitor the
 CC effect of therapy administered. The methods comprise analysing a test
 CC sample of body fluid by 2-dimensional electrophoresis to generate a 2-
 CC dimensional array of AD-associated features (AFs). The method
 CC alternatively comprises quantitatively detecting in a sample of body
 CC fluid from the subject, one or more AD-associated protein isoforms (APIs;
 CC ABR58710-ABR59184).

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 33; DB 6; Length 14;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGDLLLE 6
 DB 4 HGDLLLE 9

RESULT 4

ABR75499
 ID ABR75499 standard; peptide; 14 AA.

XX AC ABR75499;

DT 28-AUG-2003 (first entry)

XX Liver response-associated protein isoform (LRPI) peptide SEQ ID NO:84.
DE Biomarker; liver response; liver response-associated protein isoform;
KW LRPI; liver response-associated feature; LRF.
XX Synthetic.
OS WO2003038444-A2.
XX 08-MAY-2003.
XX 31-OCT-2002; 2002WO-US034847.
XX 31-OCT-2001; 2001US-0335964P.
XX (PFIZ) PEIZER PROD INC.
XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX Amacher DE, Fasulo LM, Herath HMAc, Holt GD, Stiger TR;
XX WPI; 2003-430566/40.
XX Screening, diagnosing, staging or identifying subject at risk of
XX developing, liver response, or monitoring effect of therapy on liver
XX response, by detecting Liver Response-Associated Protein Isoforms in
XX subject sample.
XX Claim 1; Page 55; 256pp; English.
XX The present invention describes a method (M1) for screening or diagnosing
XX a liver response in a subject, determining the stage or severity of a
XX liver response in a subject, identifying a subject at risk of developing
XX liver response, or monitoring the effect of therapy administered to a
XX subject having liver response, involving detecting liver response-
XX associated protein isoforms (LRPIs) (see the peptides given in ABR75395
XX to ABR75806) in a test biological sample from the subject. Alternatively,
XX screening or diagnosing a liver response in a subject, or monitoring the
XX effect of a drug or therapy administered to a subject, involves
XX contacting at least one oligonucleotide probe comprising 10 or more
XX consecutive nucleotides complementary to a nucleotide sequence encoding
XX an LRPI with RNA obtained from a biological sample from the subject or
XX with cDNA copied from the RNA, where the contacting occurs under
XX conditions that permit hybridisation of the probe to the nucleotide
XX sequence if present, detecting hybridisation, if any, between the probe
XX and the nucleotide sequence, and comparing the hybridisation, if any,
XX detected in the above step, with the hybridisation detected in a control
XX sample, or with a previously determined reference range. M1 is useful for
XX screening or diagnosing a liver response in a subject, determining the
XX stage or severity of a liver response in a subject, identifying a subject
XX at risk of developing liver response, and monitoring the effect of
XX therapy administered to a subject having liver response
XX Sequence 14 AA;
SQ
Query Match 100.0%; Score 33; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HGDLLLE 6
DB 4 HGDLLLE 9
RESULT 5
ID AAB12469 standard; peptide; 18 AA.
XX AAB12469;
AC AAB12469;
XX 25-OCT-2000 (first entry)
DT Human albumin epitope peptide #8.
XX Processing of protein or peptide samples used in mass spectrometer
DE

XX Monoclonal antibody; hybridoma cell; immunoglobulin; IgG; fusion;
KW human albumin; diabetic nephropathy; diagnosis.
XX Homo sapiens.
OS JP2000139460-A.
XX 23-MAY-2000.
XX 02-NOV-1998; 98JP-00311677.
XX 02-NOV-1998; 98JP-00311677.
XX (TOYW) TOYOTA CHUO KENKYUSHO KK.
XX (AISE) AISIN SEIKI KK.
XX WPI; 2000-433935/38.
XX Hybridoma cells for preparation of IgG monoclonal antibody capable of
XX rapid reaction with human albumin for diagnosis of diabetic nephropathy.
XX Example; Fig 1; ilpp; Japanese.
XX The present invention describes hybridoma cells used for preparing an
XX immunoglobulin G (IgG) monoclonal antibody capable of rapid reaction with
XX human albumin. The hybridoma cells are prepared by fusion of mammal
XX myeloma cells and spleen cells immunised with human serum albumin, and
XX producing IgG monoclonal antibody which rapidly react with human albumin.
XX The monoclonal antibodies can be used in the diagnosis of diabetic
XX nephropathy. The present sequence represents a human albumin epitope
XX peptide sequence which is used in an example from the present invention
XX Sequence 18 AA;
SQ
Query Match 100.0%; Score 33; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HGDLLLE 6
DB 1 HGDLLLE 6
RESULT 6
ID ABR62926 standard; peptide; 21 AA.
XX ABR62926;
AC ABR62926;
XX 04-DEC-2003 (first entry)
DT Bovine serum albumin peptide, analysed by mass spectrometry.
XX DE Serum albumin; cattle; mass spectrometry; proteomics.
XX Serun albumin; cattle; mass spectrometry; proteomics.
XX Bos taurus.
XX WO2003060524-A2.
XX 24-JUL-2003.
XX 30-DEC-2002; 2002WO-CA002024.
XX 28-DEC-2001; 2001US-0343959P.
XX (MDS-) MDS PROTEOMICS INC.
XX Duwel H, Goh T, Lebihan T;
XX WPI; 2003-618204/58.
XX Processing of protein or peptide samples used in mass spectrometer
PT

PT analysis, by reversibly immobilizing samples onto solid support,
PT transforming to solid-phase chemical unit, eluting and recovering
PT fragments from support.

XX Example 4; Page 46; 66pp; English.
XX The present sequence is a peptide fragment of bovine serum albumin (BSA).
CC The peptide was produced by chemical reduction of cysteine residues with
CC dithiothreitol, alkylation with iodoacetamide and enzymatic digestion
CC with trypsin. This is an example of the protein processing method of the
CC invention, which was designed for use in proteomics analysis by mass
CC spectrometry (MS). The general method involves: reversibly immobilising
CC protein or peptide samples onto a solid support, e.g. a cation-exchange
CC resin; subjecting the immobilised protein or peptide to solid phase-based
CC chemical and/or enzymatic modifications; eluting the resulting peptide or
CC protein derived fragments from the solid support; and recovering the
CC peptides in near-quantitative yields in a format suitable for
CC identification by routine MS technologies

XX SQ Sequence 21 AA;
Query Match 100.0%; Score 33; DB 7; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGDLLLE 6
DB 4 HGDLLLE 9

RESULT 7
ABE40958
ID ABB40958 standard; peptide; 58 AA.
XX AC ABB40958;
XX DT 04-FEB-2002 (first entry)
XX DE Peptide #8464 encoded by human foetal liver single exon probe.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000669.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human foetal liver.

XX PS Claim 27; SEQ ID NO 33593; 639pp + Sequence Listing; English.
XX CC The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human foetal liver. The

CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 58 AA;
Query Match 100.0%; Score 33; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGDLLLE 6
DB 34 HGDLLLE 39

RESULT 8
AAM34732
ID AAM34732 standard; protein; 58 AA.
XX AC AAM34732;
XX DT 17-OCT-2001 (first entry)
XX DE Peptide #8769 encoded by probe for measuring placental gene expression.
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.

XX PS Claim 27; SEQ ID NO 35001; 654pp; English.
XX CC The present invention relates to single exon nucleic acid probes (SNP:
XX see AAI31315-AAI57546). the present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders

XX SQ Sequence 58 AA;
Query Match 100.0%; Score 33; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HGDLLLE 6
DB 34 HGDLLLE 39


```

XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34925.
XX DE Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000668.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GH-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PN Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-489900/53.
XX DR
XX XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX FT gene expression in human bone marrow.
XX PT
XX PS Example 4; SEQ ID NO 34925; 658bp + Sequence Listing; English.
XX PP
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a
XX CC protein encoded by one of the probes of the invention
XX XX
XX SQ Sequence 58 AA;
XX
XX Query Match 100.0%; Score 33; DB 4; Length 58;
XX Best Local Similarity 100.0%; Pred. No. 22;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX QY 1 HGDLLLE 6
XX |||||
XX Db 34 HGDLLLE 39
XX
XX RESULT 11
XX AAM61818
XX ID AAM61818 standard; protein; 58 AA.
XX XX
XX AAM61818;
XX XX
XX DT 05-NOV-2001 (first entry)
XX XX
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33923.
XX XX
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX XX
XX PD 09-AUG-2001.
XX XX
XX 30-JAN-2001; 2001WO-US000667.
XX PF
XX XX 04-FEB-2000; 2000US-0180312P.
XX PR

```

Spatially-addressable set of single exon nucleic acid probes, used to

```

30-JAN-2001; 2001WO-US000670.
04-FEB-2000; 2000US-0180312P.
26-MAY-2000; 2000US-0207456P.
30-JUN-2000; 2000US-00608408.
03-AUG-2000; 2000US-00632366.
21-SEP-2000; 2000US-0234687P.
27-SEP-2000; 2000US-0236359P.
04-OCT-2000; 2000GB-00024263.
(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-488901/53.
Human genome-derived single exon nucleic acid probes useful for analyzing
gene expression in human cervical epithelial cells.
Claim 27; SEQ ID NO 23614; 487pp; English.
The present invention relates to human single exon nucleic acid probes
(SENP; see AII10068-AA128459). The present sequence is a peptide encoded
by one such probe. The SENPs are derived from human HeLa cells. The SENPs
can be used to produce a single exon microarray, which can be used for
measuring human gene expression in a sample derived from human cervical
epithelial cells. By measuring gene expression, the probes are therefore
useful in grading and/or staging of diseases of the cervix, notably
cervical cancer. Note: The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
Sequence 85 AA;
Query Match 100.0%; Score 33; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HGDLLS 6
DB 34 HGDLLS 39
|||||
XXXXXXXXXX
RESULT 15
ABB37889
ID ABB37889 standard; peptide; 85 AA.
AC ABB37889;
XX
DT DT DT (first entry)
XX
XX
DE Peptide #5395 encoded by human foetal liver single exon probe.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX Homo sapiens.
OS
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
(MOLE-) MOLECULAR DYNAMICS INC.
XX
XX

```

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT Gene expression in human fetal liver.
 XX
 PS Claim 27; SEQ ID NO 30524; 639pp + Sequence Listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for measuring
 CC human gene expression in a sample derived from human foetal liver. The
 CC single exon nucleic acid probes may be used for predicting, measuring and
 CC displaying gene expression in samples derived from human fetal liver. The
 CC present sequence is a peptide encoded by a single exon nucleic acid probe
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 85 AA;

 Query Match 100.0%; Score 33; DB 4; Length 85;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 HGDLLE 6
 Db 34 HGDLLE 39

 Search completed: April 19, 2004, 11:51:19
 Job time : 4.52909 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 2.52632 Seconds

(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_247_252

Perfect score: 33

Sequence: 1 HGDLL6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	33	100.0	14	9	US-09-826-290-168
2	33	100.0	14	12	US-10-601-837-46
3	33	100.0	14	15	US-10-285-394-84
4	33	100.0	14	16	US-10-264-309-77
5	33	100.0	58	9	US-09-864-761-40366
6	33	100.0	85	9	US-09-864-761-38445
7	33	100.0	99	14	US-10-029-386-29821
8	33	100.0	111	14	US-10-029-386-29303
9	33	100.0	111	14	US-10-029-386-29468
10	33	100.0	114	12	US-10-296-115-799
11	33	100.0	114	14	US-10-029-386-29057
12	33	100.0	119	14	US-10-029-386-29571
13	33	100.0	126	14	US-10-029-386-29777
14	33	100.0	244	12	US-10-425-114-57360
15	33	100.0	379	15	US-10-353-690-88

16	33	100.0	379	15	US-10-344-890-2	Sequence 2, Appli
17	33	100.0	553	12	US-10-425-114-64794	Sequence 64794, A
18	33	100.0	583	15	US-10-360-101-200	Sequence 200, App
19	33	100.0	585	9	US-09-929-552-2	Sequence 2, Appli
20	33	100.0	585	10	US-09-932-613-445	Sequence 445, App
21	33	100.0	585	10	US-09-984-010-26	Sequence 26, Appl
22	33	100.0	585	10	US-09-833-041-18	Sequence 18, Appl
23	33	100.0	585	10	US-09-833-117-18	Sequence 18, Appl
24	33	100.0	585	10	US-09-932-322-445	Sequence 445, App
25	33	100.0	585	10	US-09-832-501-18	Sequence 18, Appl
26	33	100.0	585	11	US-09-833-118-18	Sequence 18, Appl
27	33	100.0	585	11	US-09-833-245-18	Sequence 18, Appl
28	33	100.0	585	12	US-10-424-999-11	Sequence 31, Appl
29	33	100.0	585	12	US-10-425-000-31	Sequence 34, Appl
30	33	100.0	585	12	US-10-433-108-34	Sequence 5, Appli
31	33	100.0	585	13	US-10-153-064-5	Sequence 5, Appli
32	33	100.0	585	14	US-10-153-6048-5	Sequence 2, Appli
33	33	100.0	585	14	US-10-319-263-1	Sequence 2, Appli
34	33	100.0	585	14	US-10-319-263-2	Sequence 1, Appli
35	33	100.0	585	14	US-10-414-469-1	Sequence 2, Appli
36	33	100.0	585	14	US-10-414-469-2	Sequence 2, Appli
37	33	100.0	585	14	US-10-413-831-1	Sequence 2, Appli
38	33	100.0	585	14	US-10-413-831-2	Sequence 1, Appli
39	33	100.0	585	15	US-10-413-832-1	Sequence 2, Appli
40	33	100.0	585	15	US-10-413-832-2	Sequence 1, Appli
41	33	100.0	585	15	US-10-414-386-1	Sequence 2, Appli
42	33	100.0	585	15	US-10-414-386-2	Sequence 11, Appl
43	33	100.0	585	15	US-10-233-675A-11	Sequence 26, Appl
44	33	100.0	585	15	US-10-462-262-26	Sequence 7, Appli
45	33	100.0	604	10	US-09-984-010-7	

ALIGNMENTS

RESULT 1

US-09-826-290-168
Sequence 168, Application US/09826290
Patent No. US20020164668A1
GENERAL INFORMATION:
APPLICANT: Durham, L. Kathryn
APPLICANT: Friedman, David L.
APPLICANT: Herath, Herath Mudiyansele Athula Chandrasiri
APPLICANT: Kimmel, Lida H.
APPLICANT: Kimmel, Rajesh Bhikhu
APPLICANT: Parekh, David M.
APPLICANT: Potter, Christian
APPLICANT: Rohlf, B. Michael
APPLICANT: Silber, Thomas R.
APPLICANT: Stiger, Robert Reid
APPLICANT: Sunderland, P. Trey
APPLICANT: Townsend, Robert Reid
APPLICANT: White, Frost
APPLICANT: Williams, Stephen A.
TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and Treatment of
TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
FILE OF INVENTION: Alzheimer's Disease
FILE REFERENCE: 2572-1-001 N2
CURRENT APPLICATION NUMBER: US/09/826,290
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: US 60/194,504
PRIOR FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: US 60/253,647
PRIOR FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 492
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 168
LENGTH: 14
TYPE: PRT
ORGANISM: homo sapien
US-09-826-290-168

Query Match 100.0%; Score 33; DB 9; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.1;

Qy 1 HGDLE 6
|||
D'b 4 HGDLE 9

```

RESULT 5
US-09-864-761-40366
, Sequence 40366, Application US/09864761
, Patent No. US20020048763A1
, GENERAL INFORMATION:
, APPLICANT: Penn, Sharon G.
, APPLICANT: Rank, David R.
, APPLICANT: Hanzel, David K.
, APPLICANT: Chen, Wensheng
, TITLE OF INVENTION: HUMAN GENOME-DERIVED
, TITLE OF INVENTION: GENE EXPRESSION ANAL
, FILE REFERENCE: Aeomica-x-1
, CURRENT APPLICATION NUMBER: US/09/864,763
, CURRENT FILING DATE: 2001-05-23
, PRIOR APPLICATION NUMBER: US 60/180,312
, PRIOR FILING DATE: 2000-02-04
, PRIOR APPLICATION NUMBER: US 60/207,456
, PRIOR FILING DATE: 2000-05-26
, PRIOR APPLICATION NUMBER: US 09/632,366
, PRIOR FILING DATE: 2000-08-03
, PRIOR APPLICATION NUMBER: GB 24263.6
, PRIOR FILING DATE: 2000-10-04
, PRIOR APPLICATION NUMBER: US 60/236,359
, PRIOR FILING DATE: 2000-09-27
, PRIOR APPLICATION NUMBER: PCT/US01/00666
, PRIOR FILING DATE: 2001-01-30
, PRIOR APPLICATION NUMBER: PCT/US01/00667

```

PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 40366
 LENGTH: 58
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC008468.4
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.89
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
 OTHER INFORMATION: EST HUMAN HIT: A196295.1, EVALUE 2.00e-25
 OTHER INFORMATION: SWISSPROT HIT: O35865, EVALUE 4.60e+00
 US-09-864-761-40366

Query Match 100.0%; Score 33; DB 9; Length 58;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
 Db 34 HGDLL 39

RESULT 6
 US-09-864-761-38445
 Sequence 38445, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: Aecmca-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 38445
 LENGTH: 85
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC005609.1
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
 OTHER INFORMATION: SWISSPROT HIT: O35865, EVALUE 8.30e+00
 OTHER INFORMATION: EST HUMAN HIT: A196295.1, EVALUE 4.00e-41
 US-09-864-761-38445

Query Match 100.0%; Score 33; DB 9; Length 85;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
 Db 34 HGDLL 39

RESULT 7
 US-10-029-386-29821
 Sequence 29821, Application US/10029386
 Publication No. US20030194704A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharron G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 FILE REFERENCE: AECMCA-X-2
 CURRENT APPLICATION NUMBER: US/10/029,386
 CURRENT FILING DATE: 2001-12-20
 NUMBER OF SEQ ID NOS: 34288
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 29821

; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHRS.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.99
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
; OTHER INFORMATION: SWISSPROT HIT: Q9Y511, EVALUATION 5.00e-51
US-10-029-386-29468

Query Match 100.0%; Score 33; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
| | | | |
Db 42 HGDLL 47

RESULT 8

US-10-029-386-29303
; Sequence 29303, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34286
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29303
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHRS.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.65
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.49
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.41
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.47
; OTHER INFORMATION: SWISSPROT HIT: Q9UN75, EVALUATION 1.00e-58
US-10-029-386-29303

Query Match 100.0%; Score 33; DB 14; Length 111;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
| | | | |
Db 53 HGDLL 58

RESULT 9

US-10-029-386-29468
; Sequence 29468, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
US-10-029-386-29468

; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29468
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHRS.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.69
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.67
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.66
; OTHER INFORMATION: SWISSPROT HIT: Q9Y5H6, EVALUATION 2.00e-57
US-10-029-386-29468

Query Match 100.0%; Score 33; DB 14; Length 111;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
| | | | |
Db 52 HGDLL 57

RESULT 10

US-10-296-115-799
; Sequence 799, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 799
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(114)
; OTHER INFORMATION: Xaa = any amino acid or other as shown in Table 3
US-10-296-115-799

Query Match 100.0%; Score 33; DB 12; Length 114;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLL 6
| | | | |
Db 46 HGDLL 51

RESULT 11

US-10-029-386-29057
; Sequence 29057, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
US-10-029-386-29057

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29057
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR5.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: Q9Y511, EVALUE 6.00e-60
; US-10-029-386-29057

Query Match 100.0%; Score 33; DB 14; Length 114;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
|||||
DB 45 HGDLLLE 50

RESULT 12
US-10-029-386-29571
; Sequence 29571, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029.386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29571
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR5.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.71
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.85
; OTHER INFORMATION: SWISSPROT HIT: Q9Y512, EVALUE 1.00e-58
; US-10-029-386-29571

Query Match 100.0%; Score 33; DB 14; Length 119;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
|||||
DB 60 HGDLLLE 65

RESULT 13
US-10-029-386-29777
; Sequence 29777, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029.386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29777
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR5.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.82
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.67
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: Q9UN75, EVALUE 6.00e-63
; US-10-029-386-29777

Query Match 100.0%; Score 33; DB 14; Length 126;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
|||||
DB 64 HGDLLLE 69

RESULT 14
US-10-425-114-57360
; Sequence 57360, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57360
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17312D02_FLI.pep
; US-10-425-114-57360

Query Match 100.0%; Score 33; DB 12; Length 244;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
|||||
DB 231 HGDLLLE 236

RESULT 15
US-10-353-690-88
; Sequence 88, Application US/10353690
; Publication No. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary

```

; APPLICANT: Staaliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodrique-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: Cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2524, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
; FILE REFERENCE: MPI02-018P1RNMIM
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-353-690-88

Query Match          100.0%; Score 33; DB 15; Length 379;
Best Local Similarity 100.0%; Pred.No.1.2e+02;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLE 6
Db      93 HGDLE 98

Search completed: April 19, 2004, 12:54:58
Job time : 2.52632 secs
```

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	47	100.0	584	1	US-08-448-196A-17	Sequence 7, Appli	
2	47	100.0	585	1	US-08-453-799-14	Sequence 14, Appli	
3	47	100.0	585	1	US-08-448-136A-3	Sequence 3, Appli	
4	47	100.0	585	2	US-08-984-176-1	Sequence 1, Appli	
5	47	100.0	585	2	US-08-703-572-2	Sequence 2, Appli	
6	47	100.0	585	3	US-08-769-746-5	Sequence 2, Appli	
7	47	100.0	585	4	US-10-153-064-5	Sequence 5, Appli	
8	47	100.0	609	1	US-08-222-613-3	Sequence 3, Appli	
9	47	100.0	609	1	US-08-433-037-4	Sequence 4, Appli	
10	47	100.0	609	4	US-08-897-956A-2	Sequence 2, Appli	
11	47	100.0	609	4	US-10-153-064-7	Sequence 7, Appli	
12	47	100.0	609	4	US-09-976-594-977	Sequence 977, App	
13	47	100.0	609	5	PCR-US95-04075-3	Sequence 3, Appli	
14	47	100.0	610	2	US-08-797-683-2	Sequence 2, Appli	
15	47	100.0	610	4	US-09-984-186-2	Sequence 2, Appli	
16	47	100.0	651	4	US-10-153-064-133	Sequence 133, App	
17	47	100.0	652	4	US-10-153-064-96	Sequence 96, Appli	
18	47	100.0	652	4	US-10-153-064-99	Sequence 99, Appli	
19	47	100.0	652	4	US-10-153-064-105	Sequence 105, App	
20	47	100.0	652	4	US-10-153-064-132	Sequence 132, App	
21	47	100.0	653	4	US-10-153-064-131	Sequence 131, App	
22	47	100.0	656	4	US-10-153-064-130	Sequence 130, App	
23	47	100.0	660	4	US-10-153-064-90	Sequence 90, Appli	
24	47	100.0	660	4	US-10-153-064-93	Sequence 93, Appli	
25	47	100.0	668	4	US-10-153-064-102	Sequence 102, App	
26	47	100.0	676	4	US-10-153-064-95	Sequence 95, Appli	
27	47	100.0	676	4	US-10-153-064-98	Sequence 98, Appli	

Db 92 AKQEPERNE 100

RESULT 2

US-08-153-799-14

; Sequence 14, Application US/08153799

; Patent No. 5766883

GENERAL INFORMATION:

APPLICANT: Ballance, David J

APPLICANT: Goodey, Andrew R

TITLE OF INVENTION: Polypeptides

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: R Hain Swope, BOC Health Care Inc

STREET: 100 Mountain Avenue

CITY: Murray Hill

STATE: New Jersey

COUNTRY: USA

ZIP: 07974

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/153.799

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/847975

FILING DATE: 06-MAR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 8909916.2

FILING DATE: 29-APR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB90/00650

FILING DATE: 26-APR-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/775952

FILING DATE: 29-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Swope, R Hain

REGISTRATION NUMBER: 24864

REFERENCE/DOCKET NUMBER: 92H832

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 665 2400

TELEFAX: (908) 771 6159

TELEX: 219484

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 585 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: Region

LOCATION: 369..419

OTHER INFORMATION: /note= "Alternative C-termini of

OTHER INFORMATION: HSA(1-n)

FEATURE:

NAME/KEY: Region

LOCATION: 1..585

OTHER INFORMATION: /note= "Amino acid sequence of

OTHER INFORMATION: natural HSA"

US-08-153-799-14

Query Match 100.0%; Score 47; DB 1; Length 585;

Best Local Similarity 100.0%; Fred. No. 0.36;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 92 AKQEPERNE 100

RESULT 3

US-08-448-196A-3

; Sequence 3, Application US/08448196A

; Patent No. 5780594

GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C.

TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

TITLE OF INVENTION: RELATED PROTEINS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: NASA

STREET: MARSHALL SPACE FLIGHT CENTER

CITY: HUNTSVILLE

STATE: ALABAMA

COUNTRY: USA

ZIP: 35812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/448.196A

FILING DATE: 23-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: BROAD JR., ROBERT L.

REGISTRATION NUMBER: 18,757

REFERENCE/DOCKET NUMBER: XX/MFS-28402-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 205-544-0021

TELEFAX: 205-544-0258

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 585 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

US-08-448-196A-3

Query Match 100.0%; Score 47; DB 1; Length 585;

Best Local Similarity 100.0%; Fred. No. 0.36;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 92 AKQEPERNE 100

RESULT 4

US-08-984-176-1

; Sequence 1, Application US/08984176

; Patent No. 5948609

GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C

APPLICANT: HO, JOSEPH X

APPLICANT: RUKER, FLORIAN

TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT

TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984.176

CURRENT APPLICATION NUMBER: US/08/984.176

CURRENT FILING DATE: 1997-12-03

NUMBER OF SEQ ID NOS: 1

SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1

Query Match      100.0%; Score 47; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 5
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match      100.0%; Score 47; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 6
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
```

```
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match      100.0%; Score 47; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 7
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match      100.0%; Score 47; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPERNE 9
Db      92 AKQEPERNE 100
```

RESULT 8
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/ARC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match 100.0%; Score 47; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
DB 116 AKQEPERNE 124

RESULT 9
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
; APPLICANT: Barr, Kathryn A.
; APPLICANT: Brierley, Russell A.
; APPLICANT: Thill, Gregory P.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
; TITLE OF INVENTION: PICHIA PASTORIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/433,037
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9108Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 100.0%; Score 47; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
DB 116 AKQEPERNE 124

RESULT 10
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion Polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
DB 116 AKQEPERNE 124

RESULT 11
US-10-153-064-7
; Sequence 7, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match      100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
   |||||
Db 116 AKQEPERNE 124

RESULT 12
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match      100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
   |||||
Db 116 AKQEPERNE 124

RESULT 13
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04075
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 3:

; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match      100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
   |||||
Db 116 AKQEPERNE 124

RESULT 14
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3043
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-36,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-2

Query Match      100.0%; Score 47; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match      100.0%; Score 47; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
   |||||
Db 116 AKQEPERNE 124

RESULT 14
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3043
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-36,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-2

Query Match      100.0%; Score 47; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Mon Apr 19 13:27:30 2004

QY 1 AKOEPERNE 9
 |||||
 Db 116 AKOEPERNE 124

RESULT 15

US-09-984-186-2
 ; Sequence 2, Application US/09984186
 ; Patent No. 6886179

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard
 Fournier, Alain
 Guitton, Jean-Dominique
 Jung, Gerard
 Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, 3043
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: System 7.1
 SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/984,186
 FILING DATE: 29-Oct-2001
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/797,689
 FILING DATE: 31-JAN-1997
 APPLICATION NUMBER: US 08/256,927
 FILING DATE: 28-JUL-1994
 APPLICATION NUMBER: FR 92/01064
 FILING DATE: 31-JAN-1992
 APPLICATION NUMBER: PCT/FR93/00085
 FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Smith Ph.D., Julie K.
 REGISTRATION NUMBER: P-38,619
 REFERENCE/DOCKET NUMBER: ST92006-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 454-3839
 TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-984-186-2

Query Match 100.0%; Score 47; DB 4; Length 610;
 Best Local Similarity 100.0%; Pred No. 0.39;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKOEPERNE 9
 |||||
 Db 116 AKOEPERNE 124

Search completed: April 19, 2004, 12:05:19
 Job time : 1.3795 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 0.853186 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_170_176

Perfect score: 32

Sequence: 1 QAADKAA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*

- 1: _pir1.*
- 2: _pir2.*
- 3: _pir3.*
- 4: _pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	32	100.0	600	A47391	serum albumin prec
2	32	100.0	605	1 ABPGS	serum albumin prec
3	32	100.0	609	1 ABHUS	serum albumin prec
4	29	90.6	96	2 AH3273	hypothetical cyto
5	29	90.6	128	2 JCI273	ribosomal protein
6	29	90.6	133	2 T11792	ribosomal protein
7	29	90.6	311	2 AG3133	2-hydroxyacid dehy
8	29	90.6	311	2 E96154	hypothetical prote
9	29	90.6	315	2 D95295	probable D-isomer
10	29	90.6	387	2 E70716	probable succinyl-
11	29	90.6	608	2 S57632	serum albumin prec
12	29	90.6	906	2 AD3267	protein translocas
13	28	87.5	260	2 C83362	hypothetical prote
14	28	87.5	293	2 B87692	carbonic anhydrase
15	28	87.5	409	2 A43256	mobilization prote
16	28	87.5	462	2 A86717	conserved hypothet
17	28	87.5	523	2 B38145	invariant surface
18	28	87.5	570	2 AD0223	flagellar M-ring p
19	28	87.5	608	2 A46312	gag polypeptide -
20	28	87.5	619	2 A41971	surface protein ps
21	28	87.5	619	2 A97887	surface protein ps
22	28	87.5	919	2 T37062	probable transcrip
23	28	87.5	1061	2 D69799	cyclochrome P450 /
24	28	87.5	1337	2 T30291	dextranase - Strept
25	27	84.4	60	2 H95354	protein [imported
26	27	84.4	130	1 R35C8	ribosomal protein
27	27	84.4	130	2 C91150	30S ribosomal subu
28	27	84.4	130	2 G85995	30S ribosomal subu
29	27	84.4	130	2 AG1007	30S ribosomal chai

outer membrane pro
outer membrane pro
histone-like prote
hypothetical prote
outer membrane pro
cationic 19 kDa ou
MHC cell surface a
hypothetical prote
hypothetical prote
sorbose-permease P
probable sorbose P
phosphotransferase
phosphotransferase
mannose-specific P
PTS enzyme IIC, ma
phosphotransferase

30 27 84.4 161 1 DNEC17
31 27 84.4 161 1 S09104
32 27 84.4 161 2 D90651
33 27 84.4 161 2 D85502
34 27 84.4 161 2 AC0530
35 27 84.4 165 2 AD0129
36 27 84.4 181 2 I79640
37 27 84.4 239 2 AE1317
38 27 84.4 243 2 AE1689
39 27 84.4 265 2 H91253
40 27 84.4 265 2 D86094
41 27 84.4 266 1 MQECMP
42 27 84.4 266 2 S0189
43 27 84.4 266 2 H90944
44 27 84.4 266 2 C85793
45 27 84.4 266 2 AG0726

ALIGNMENTS

RESULT 1

A47391
serum albumin precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A47391
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwalet, J.; Putnam, F
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilin
A:Reference number: A47391; MUID:93211971; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-600 <WAT>
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PId:G342295
A:Experimental source: liver
A:Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIF:128281)
C:Superfamily: serum albumin; serum albumin repeat homology
F:21-194/Domain: serum albumin repeat homology <SA1>
F:213-386/Domain: serum albumin repeat homology <SA2>
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 32; DB 2; Length 600;

Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7
|||
Db 186 QAADKAA 192

RESULT 2

ABPGS
serum albumin precursor - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S01382; A61006
R:Weinstock, J.; Baldwin, G. S.
Nucleic Acids Res. 16, 9045, 1988
A:Title: Nucleotide sequence of porcine liver albumin.
A:Reference number: S01382; MUID:89016582; PMID:3174440
A:Accession: S01382
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-605 <WEI>
A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798
F:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
J. Bone Miner. Res. 4, 235-241, 1989
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of minera
A:Reference number: A61006; MUID:89269769; PMID:2728927
A:Accession: A61006

A:Molecule type: protein
A:Residues: 23-51,'X',53-54,'XXXGY',146,'E',148,'E',150-151,'XVN',155 <LIM>
A:Experimental source: dental enamel
A:Note: albumin and other serum proteins are also found in bone
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membra
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
F:17-22/Domain: propeptide #status predicted <PRO>
F:23-605/Product: serum albumin #status predicted <MAT>
F:27-199/Domain: serum albumin repeat homology <SA1>
F:218-391/Domain: serum albumin repeat homology <SA2>
F:410-589/Domain: serum albumin repeat homology <SA3>
F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4
F:266/Binding site: bilirubin (lys) #status predicted
Query Match 100.0%; Score 32; DB 1; Length 605;
Beat Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQXAA 7
Db 191 QAAQXAA 197
RESULT 3
ABUS
serum albumin precursor [validated] - human
N/Alternate names: preproalbumin
N/Contains: kinetensin
C/Species: Homo sapiens (man)
C/Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000
C/Accession: A93743; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S39
R/Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur
Nucleic Acids Res. 9, 6103-6114, 1981
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli
A:Reference number: A93743; MUID:82081982; PMID:6171778
A:Accession: A93743
A:Molecule type: mRNA
A:Residues: 1-419,'K',421-609 <LAW>
A:Cross-references: EMBL:V00495; GB:J000078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA23
R/Dugaiczkyk, A.; Law, S.W.; Dennison, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A:Reference number: A93936; MUID:82105994; PMID:6275391
A:Accession: A93936
A:Molecule type: mRNA
A:Residues: 1-120,'G',122-609 <DUG>
A:Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590
R/Drano, V.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A:Reference number: I39427; MUID:86140099; PMID:2419329
A:Accession: I39427
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-26 <URA>
A:Cross-references: GB:M13075; NID:G178330; PIDN:AAA51688.1; PID:G553173
R/Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.
A:Reference number: I59286; MUID:94181575; PMID:8134387
A:Accession: I59286
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 282-290,'KSFEDIQ' <WAT>
A:Cross-references: GB:869192; NID:G546032; PIDN:AB30282.1; PID:G546033
A:Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R/Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-
A:Reference number: I59313; MUID:94294404; PMID:8022807

A/Accession: I59313
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 589-590,'ALPRVKNLLQVQLP' <MAD>
A:Cross-references: GB:S70799; NID:G547231; PIDN:AB31177.1; PID:G547232
A:Note: this frame-shift variant is designated albumin Bazzano; four additional variant:
R/Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A:Reference number: G08292
A:Accession: G01747
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-120,'G',122-455 <MEN>
A:Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431
R/Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex
A:Reference number: S55314; MUID:95275251; PMID:7755581
A:Accession: S55314
A:Molecule type: protein
A:Residues: 19-27 <LED>
R/Meloun, B.; Moravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A:Title: Complete amino acid sequence of human serum albumin.
A:Reference number: A91420; MUID:76187807; PMID:1225573
A:Accession: A91420
A:Molecule type: protein
A:Residues: 23-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',395
R/Roehr, U.; Spitteller, G.; Iripier, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A:Title: Isolation and structure elucidation of middle-molecular weight peptides from u
A:Reference number: S06422
A:Note: this paper is in German, with an English abstract
A:Accession: S06422
A:Molecule type: protein
A:Residues: 25-48 <ROE>
R/Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A:Title: Mass spectrometric identification of modifications to human serum albumin treat
A:Reference number: S36882; MUID:93384321; PMID:8373198
A:Accession: S36882
A:Molecule type: protein
A:Residues: 45-67;141-150;311-337;469-490;570-581 <FIN>
R/Kausler, E.; Spitteller, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelm
A:Reference number: S17599; MUID:92126241; PMID:1772598
A:Accession: S17599
A:Molecule type: protein
A:Residues: 25-54;354-357;431-447 <KAU>
A:Note: 49-Leu was also found
R/Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid protea
A:Reference number: A45800; MUID:89341406; PMID:2474609
A:Accession: A45800
A:Molecule type: protein
A:Residues: 166-173,'L' <MOG>
R/Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; W
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A:Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tr
A:Reference number: A03239; MUID:86242180; PMID:3087352
A:Accession: A03239
A:Molecule type: protein
A:Residues: 166-173,'L' <MOG>
R/Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins,
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A:Title: Mutations in genetic variants of human serum albumin found in Italy.
A:Reference number: A38255; MUID:91062352; PMID:2247440
A:Accession: A38255
A:Molecule type: protein
A:Residues: 76-111 <GAL1>
A:Accession: B38255

A:Molecule type: protein
A:Residues: 82-105,'K',107-110 <GAL2>
A>Note: this variant is designated albumin Vibo Valentia
A:Accession: A38255
A:Molecule type: protein
A:Residues: 76-83,'K',85-106 <GAL3>
A>Note: this variant is designated albumin Torino
R:Minichiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A:Title: The structural characterization and bilirubin-binding properties of albumin Her
A:Reference number: S33298; MUID:93292504; PMID:8513793
A:Accession: S33298
A:Molecule type: protein
A:Residues: 255-263,'E',265-281 <MIN1>
A>Note: this variant is designated albumin Herborn
R:Minichiotti, L.; Galliano, M.; Stopponi, M.C.; Perri, G.; Crespeau, H.; Rochu, D.; Porta,
Biochim. Biophys. Acta 1119, 232-238, 1992
A:Title: Two albumins with identical electrophoretic mobility are produced by differ
A:Reference number: S21078; MUID:92190239; PMID:1347703
A:Accession: S21078
A:Molecule type: protein
A:Residues: 354-356,'K',358-378 <MIN2>
A>Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,
R:He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A:Title: Atomic structure and chemistry of human serum albumin.
A:Reference number: A46756; MUID:92334427; PMID:1630489
A:Contents: annotation; X-ray crystallography, 2.8 angstroms
R:Brown, J.R.; Shockley, P.; Behrens, P.Q.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,
A:Reference number: A94442
A:Contents: annotation; three-dimensional structure and disulfide bonds
R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A:Title: Disulfide bonds in human serum albumin.
A:Reference number: A30930
A:Contents: annotation; disulfide bonds
R:Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A:Reference number: A30299; MUID:78186630; PMID:656055
A:Contents: annotation; bilirubin-binding site
R:Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20,
A:Title: Serum albumin: conformation and active sites.
A:Reference number: A94408
A:Contents: annotation; binding sites
R:Harper, M.E.; Dugaiczky, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes
A:Reference number: A90028; MUID:83279982; PMID:6192711
A:Contents: annotation; gene position
R:Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.
A:Reference number: A46755; MUID:76257808; PMID:955075
A:Contents: annotation
A>Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid
R:Bohney, J.P.; Fonda, M.L.; Feidhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphat
A:Reference number: A56294; MUID:92183881; PMID:1544460
A:Contents: annotation
A>Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p
ase activity
C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
C:Comment: A large number of variants of human serum albumin have been described.
C:Genetics:
A:Gene: GDB:ALB
A:Cross-references: GDB:118990; OMIM:103600
A:Map position: 4q11-4q13
C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox
F:1-19/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-609/Product: serum albumin #status experimental <MPT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:166-174/Product: kinetensin #status experimental <KIP>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:27/Binding site: copper (His) #status predicted
F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental
Query Match 100.0%; Score 32; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAADKAA 7
DB 194 QAADKAA 200
RESULT 4
AH3273
Hypothetical cytosolic protein BMS10173 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AH3273
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, C.; Mujar, C.; Los, T.; Ivanova,
Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes;
Proc.Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella meliten
A:Reference number: AD3252; PMID:11756688
A:Accession: AH3273
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL51355.1; PID:gi7982054; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Map position: 1
Query Match 90.6%; Score 29; DB 2; Length 96;
Best Local Similarity 85.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAADKAA 7
DB 56 EAADKAA 62
RESULT 5
JC1273
ribosomal protein L7/L12 - Streptomyces antibioticus
C:Species: Streptomyces antibioticus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
C:Accession: JC1273
R:Parra, F.; Blanco, G.; Alonso, J.M.M.; Balbin, M.; Mendez, C.; Salas, J.A.
Gene 118, 127-129, 1992
A:Title: Cloning and sequence of a gene encoding the L7/L12 ribosomal protein equivalent
A:Reference number: JC1273; MUID:92380478; PMID:1511874
A:Accession: JC1273
A:Molecule type: DNA
A:Residues: 1-128 <PAR>
A:Cross-references: GB:M89911; NID:gi53436; PIDN:AAA26811.1; PID:gi53438
C:Superfamily: Escherichia coli ribosomal protein L12
C:Keywords: protein biosynthesis; ribosome
Query Match 90.6%; Score 29; DB 2; Length 128;
Best Local Similarity 85.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAADKAA 7

```

Db      109 EAADKAA 115
      :|||||
      |
RESULT 6
ribosomal protein L12 - Streptomyces virginiae
T11792
C:Species: Streptomyces virginiae
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C:Accession: T11792
R:Yamada, Y.
submitted to the EMBL Data Library, May 1995
A:Reference number: Z17345
A:Accession: T11792
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-133 <YAM>
A:Cross-references: EMBL:D50624
C:Genetics:
A:Gene: rplL
C:Superfamily: Escherichia coli ribosomal protein L12
C:Keywords: ribosome

      Query Match      90.6%; Score 29; DB 2; Length 133;
      Best Local Similarity 85.7%; Pred. No. 24;
      Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

      QY      1 QAADKAA 7
      :|||||
      |
      Db      114 EAADKAA 120

RESULT 7
2-hydroxyacid dehydrogenase Atu4691 [imported] - Agrobacterium tumefaciens (strain C58,
AG3133
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AG3133
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG3133
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-311 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAL45485.1; PID:gl7743192; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4691
A:Map position: linear chromosome
C:Superfamily: phosphoglycerate dehydrogenase

      Query Match      90.6%; Score 29; DB 2; Length 311;
      Best Local Similarity 85.7%; Pred. No. 54;
      Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

      QY      1 QAADKAA 7
      :|||||
      |
      Db      30 EAADKAA 36

RESULT 8
hypothetical protein AGR_L379 [imported] - Agrobacterium tumefaciens (strain C58, Cerec
E98154
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C:Accession: E98154
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,

```

```

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: E98154
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-311 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK88759.1; PID:gl5158503; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L379
A:Map position: linear chromosome
C:Superfamily: phosphoglycerate dehydrogenase

      Query Match      90.6%; Score 29; DB 2; Length 311;
      Best Local Similarity 85.7%; Pred. No. 54;
      Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

      QY      1 QAADKAA 7
      :|||||
      |
      Db      30 EAADKAA 36

RESULT 9
D95295
probable D-isomer specific 2-hydroxyacid [imported] - Sinorhizobium meliloti (strain 10
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: D95295
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9889, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: D95295
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-315 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK64926.1; PID:gl4523348; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMA0510
A:Genome: plasmid
C:Superfamily: phosphoglycerate dehydrogenase

      Query Match      90.6%; Score 29; DB 2; Length 315;
      Best Local Similarity 85.7%; Pred. No. 55;
      Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

      QY      1 QAADKAA 7
      :|||||
      |
      Db      31 EAADKAA 37

RESULT 10
E70716
probable succinyl-coa synthetase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: E70716
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

```

A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70716
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-387 <COL>
A;Cross-references: GB:Z79700; GB:AL123456; NID:g3261628; PIDN:CAB01999.1; PID:gl524211
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: succ
C;Superfamily: succinate-CoA ligase (ADP-forming) beta chain

Query Match 90.6%; Score 29; DB 2; Length 387;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
:|||||

DB 376 EAADKAA 382

RESULT 11

SS7632
serum albumin precursor - cat
C;Species: Felis silvestris catus (domestic cat)
C;Date: 18-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
C;Accession: JC4660; S57632
R;Hilger, C.; Grigioni, F.; Hentges, F.
Gene 169, 295-296, 1996
A;Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A;Reference number: JC4660; MUID:96194824; PMID:8647469
A;Accession: JC4660
A;Molecule type: mRNA
A;Residues: 1-608 <H12>
A;Cross-references: EMBL:X84842; NID:g886484; PIDN:CAA59279.1; PID:g886485
A;Experimental source: liver
C;Comment: This protein is the major protein component in plasma. It functions as a multiein has 35 conserved cysteine residues.
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: liver; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRP>
F;25-608/Product: serum albumin #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>

Query Match 90.6%; Score 29; DB 2; Length 608;
Best Local Similarity 85.7%; Pred. No. 16+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
:|||||

DB 194 EAADKAA 200

RESULT 12

AD3267
protein translocase, chain secA [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AD3267
R;DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Ios, T.; Ivanova, M.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3262; PMID:11756688
A;Accession: AD3267
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-905 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL51303.1; PID:gl7981998; GSPDB:GN00190

Query Match 90.6%; Score 29; DB 2; Length 906;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
:|||||

DB 736 EAADKAA 742

RESULT 13

C83362
hypothetical protein PA2260 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: C83362
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: C83362
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-260 <STO>
A;Cross-references: GB:AE004652; GB:AE004827; PIDN:AA05648.1; GSPDB:GN00:
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2260

Query Match 87.5%; Score 28; DB 2; Length 260;
Best Local Similarity 85.7%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
:|||||

DB 170 QAADKAA 176

RESULT 14

B87692
carbonic anhydrase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: B87692
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4138-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87692
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-233 <STO>
A;Cross-references: GB:AE005673; NID:gl3425312; PIDN:AAK25534.1; GSPDB:GN00148
C;Genetics:
A;Gene: CC3572

Query Match 87.5%; Score 28; DB 2; Length 293;
Best Local Similarity 85.7%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAADKAA 7
:|||||

DB 64 QAADKAA 70

RESULT 15
A43256
mobilization protein mobA - Thiobacillus ferrooxidans plasmid pTF-PC2
C:Species: Thiobacillus ferrooxidans
C>Date: 10-Jun-1993 #sequence_revision 01-Dec-1995 #text_change 08-Oct-1999
C/Accession: A43256; S27622
R/Rohrer, J.; Rawlings, D.E.
J. Bacteriol. 174, 6230-6237, 1992
A/Title: Sequence analysis and characterization of the mobilization region of a broad-host-range plasmid from Thiobacillus ferrooxidans
A/Reference number: A43256; MUID:93015664; PMID:1400173
A/Accession: A43256
A/Status: Preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-409 <ROH>
A/Cross-references: EMBL:MS717; NID:G154659; PIDN:AAA27389.1; PID:G154660
A/Note: sequence extracted from NCBI backbone (NCBIP:115305)
C/Genetics:
A/Gene: mobA
A/Genome: plasmid

Query Match 87.5%; Score 28; DB 2; Length 409;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAAAKAA 7
|||
Db 306 QAAEKAA 312

Search completed: April 19, 2004, 12:02:25
Job time : 1.85319 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 5.29363 Seconds
(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_280_288

Perfect score: 47

Sequence: 1 EXPLEKSH 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	100.0	112	AA02636	Human pol
2	47	100.0	114	AA02636	Human pro
3	47	100.0	123	AA023861	abcEST en
4	47	100.0	126	AA004435	Human pol
5	47	100.0	188	AA083948	Yeast cod
6	47	100.0	293	AA033082	Novel hum
7	47	100.0	303	AA014178	Human ser
8	47	100.0	373	AA090387	N-termina
9	47	100.0	388	AA090389	N-termina
10	47	100.0	389	AA090390	N-termina
11	47	100.0	390	AA090391	N-termina
12	47	100.0	407	AA090392	N-termina
13	47	100.0	463	AA014179	Human ser
14	47	100.0	500	AD032019	Heterolog
15	47	100.0	550	AA029877	Novel hum
16	47	100.0	584	AB072381	Mature hu
17	47	100.0	585	AA093344	Sequence
18	47	100.0	585	AA090388	Mature hu
19	47	100.0	585	AA091422	Human nor
20	47	100.0	585	AA05318	Human ser
21	47	100.0	585	AA08457	Human ser
22	47	100.0	585	AA02607	Human ser
23	47	100.0	585	AA026362	Synthetic
24	47	100.0	585	AA020029	Human ser
25	47	100.0	585	AA080301	Human ser

26	47	100.0	585	AA020111	HSA prote
27	47	100.0	585	AA059841	Mature pr
28	47	100.0	585	AA084873	Amino aci
29	47	100.0	585	AA083946	Yeast cod
30	47	100.0	585	AA052567	Mature hu
31	47	100.0	585	AA012417	Human alb
32	47	100.0	585	AA012403	Human alb
33	47	100.0	585	AA013129	Human alb
34	47	100.0	585	AA013135	Human alb
35	47	100.0	585	AA013111	Human alb
36	47	100.0	585	AA013399	Human alb
37	47	100.0	585	AA079006	Human mat
38	47	100.0	585	AA08578	Human ser
39	47	100.0	585	AA075220	Mature fo
40	47	100.0	585	AA070986	B lymphoc
41	47	100.0	585	AB063321	Human ser
42	47	100.0	585	AB033847	Human B L
43	47	100.0	585	AB071291	Glycosyla
44	47	100.0	585	AB055695	Human alb
45	47	100.0	585	AA030936	Human alb

ALIGNMENTS

RESULT 1
AA002636
ID AA002636 standard; protein; 112 AA.

XX AA002636;
AC AA002636;

XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 16528.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.

XX N-PSDB; AA02557.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 16528; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA079941-AA093841) and
XX the encoded proteins (AA00010-AA013910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The

XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ

Sequence 112 AA;

Query Match 100.0%; Score 47; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKPLLEKSH 9

DB 3 EKPLLEKSH 11

RESULT 2

AAAM25284
ID AAAM25284 standard; protein; 114 AA.

AC AAAM25284;

DT 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:799.

XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnary; antitumor; osteopathic; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.

XX Homo sapiens.

XX WO200153455-A2.

XX 26-JUL-2001.

XX 22-DEC-2000; 2000WO-US035017.

XX 23-DEC-1999; 98US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457603/49.

XX DR N-PSDB; AAH99225.

XX Isolated human polynucleotides encoding polypeptides, useful for the
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.

XX Claim 20; Page 186; 1217pp; English.

XX AAH99166 to AAH99904 encode the human proteins given in AAAM25225 to
CC AAAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnary;
CC antitumor; osteopathic; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;

CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antitense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders

XX Sequence 114 AA;

Query Match 100.0%; Score 47; DB 4; Length 114;

Best Local Similarity 100.0%; Pred. No. 0.11;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EKPLLEKSH 9

DB 79 EKPLLEKSH 87

RESULT 3

AAAM23861

ID AAAM23861 standard; protein; 123 AA.

AC AAAM23861;

DT 12-OCT-2001 (first entry)

DE abCEST encoded protein SEQ ID NO: 1386.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;

XX tomato; monkey; dog; sea urchin; expressed sequence tag; EST;

XX diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
XX gene therapy; nutrition.

XX Unidentified.

XX WO200154477-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US002687.

XX 25-JAN-2000; 2000US-00491404.

XX 17-JUL-2000; 2000US-00617746.

XX 03-AUG-2000; 2000US-00631451.

XX 15-SEP-2000; 2000US-00683870.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;

XX Cao Y, Drmanac RA, Zhang J, Werhman T;

XX WPI; 2001-476164/51.

XX DR N-PSDB; AAH98520.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising
XX antibodies and research use.

XX Claim 20; Page 975; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a

CC protein of the invention
XX
SQ Sequence 123 AA;

Query Match 100.0%; Score 47; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 77 EKPLLEKSH 85

RESULT 4
AAO04435
ID AAC004435 standard; protein; 126 AA.
XX
AC AAC04435;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 18327.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
FN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI84366.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 20; SEQ ID NO 18327; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 126 AA;

Query Match 100.0%; Score 47; DB 4; Length 126;
Best Local Similarity 100.0%; Pred. No. 0.13; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 77 EKPLLEKSH 85

Db 118 EKPLLEKSH 126

RESULT 5
AAV83948
ID AAV83948 standard; protein; 188 AA.
XX
AC AAV83948;
XX
DT 28-JUL-2000 (first entry)
XX
DE Yeast codon-biased recombinant HSA protein fragment HSA-II.
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
KW overlapping oligonucleotide; expression vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CNI239103-A.
XX
PD 22-DEC-1999.
XX
PF 17-JUN-1998; 98CN-00102506.
XX
PR 17-JUN-1998; 98CN-00102506.
XX
PA (HAJ-) HAJI BIOENGINEERING CO LTD.
XX
PI Li S, Lu D;
XX
DR WPI; 2000-351198/31.
DR N-PSDB; AAA10093.
XX
PT Process for preparing recombinant human serum albumin comprising yeast
PT biased sex codons - uses a recombinant DNA technique.
XX
PS Example 1; Fig 5; 44pp; Chinese.
XX
CC The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AA10091) was generated as
CC three synthetic fragments (AA10092-AA10094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the sequence of the HSA fragment HSA-II encoded by the human gene with a
CC yeast codon bias. The invention also covers a recombinant expression
CC vector, yeast host cells carrying the recombinant expression vector and
CC the process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
SQ Sequence 188 AA;

Query Match 100.0%; Score 47; DB 3; Length 188;
Best Local Similarity 100.0%; Pred. No. 0.19; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 96 EKPLLEKSH 104

RESULT 6
AAU33082
ID AAU33082 standard; protein; 293 AA.
XX
AC AAU33082;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3573.
XX
KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX Homo sapiens.
XX WO200179449-A2.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US008656.
XX
XX 18-APR-2000; 2000US-00552929.
XX
XX 26-JAN-2001; 2001US-00770160.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX
XX Claim 20; Page 705; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX and/or nerve tissue growth or regeneration; immune suppression and/or
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX AAU29510-AAU33304 represent the amino acid sequences of novel human
XX secreted proteins of the invention
XX
XX Sequence 293 AA;
XX
Query Match 100.0%; Score 47; DB 4; Length 293;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EKPLLEKSH 9
DB 134 EKPLLEKSH 142
XX
RESULT 7
AAR14178
ID AAR14178 standard; protein; 303 AA.
XX
XX AAR14178;
XX
XX 19-DEC-1991 (first entry)
XX
XX Human serum albumin lacking C-terminal fragment.
XX
XX HSA; fusion protein; drug.
XX
XX Homo sapiens.
XX
XX JP03201987-A.
XX
XX 03-SEP-1991.
XX
XX 29-DEC-1989; 89JP-00344701.
XX

XX 29-DEC-1989; 89JP-00344701.
XX (TOFU) TONEN CORP.
XX
XX WPI; 1991-300976/41.
XX
XX Human serum albumin fragment - where C-terminal of human serum albumin is
XX lacking and which can be combined with various drugs.
XX
XX Claim 1; Page 1; 23pp; Japanese.
XX
XX This sequence corresponds to amino acids 1 to 303 of mature human serum
XX albumin. The fragment lacking the C-terminal sequence can form part of a
XX fusion protein, for example with drugs. (This sequence is taken from the
XX full-length HSA sequence in EP-330451). See also AAR14179
XX
XX Sequence 303 AA;
XX
Query Match 100.0%; Score 47; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EKPLLEKSH 9
DB 280 EKPLLEKSH 288
XX
RESULT 8
AAP90387
ID AAP90387 standard; protein; 373 AA.
XX
XX AAP90387;
XX
XX 24-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 01-NOV-1989 (first entry)
XX
XX N-terminal of human serum albumin polypeptide.
XX
XX Human serum albumin polypeptide; plasma expanders.
XX
XX Homo sapiens; (Human).
XX
XX EP322094-A.
XX
XX 28-JUN-1989.
XX
XX 25-OCT-1988; 88EP-00310000.
XX
XX 30-OCT-1987; 87GB-00025529.
XX
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX WPI; 1989-186464/26.
XX
XX New N-terminal fragments of human serum albumin - esp. useful as blood
XX plasma expanders.
XX
XX Claim 2; Page 9; 20pp; English.
XX
XX N-terminal portion of human serum albumin. Used as plasma expanders, or
XX as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-
XX MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
XX field)
XX
XX Sequence 373 AA;
XX
Query Match 100.0%; Score 47; DB 1; Length 373;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKPLEKSH 9
Db 280 EKPLEKSH 288

RESULT 9
AAP90389
ID AAP90389 standard; protein; 388 AA.
XX AC AAP90389;
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX N-terminal human serum albumin polypeptide.
XX N-terminal human serum albumin polypeptide; plasma expanders.
XX Homo sapiens; (Human).
XX EP322094-A.
XX 28-JUN-1989.
XX 25-OCT-1988; 88EP-00310000.
XX 30-OCT-1987; 87GB-00025529.
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX WPI; 1989-186464/26.
XX New N-terminal fragments of human serum albumin - esp. useful as blood plasma expanders.
XX Claim 2; Page 9; 20pp; English.
XX N-terminal portion of human serum albumin. Used to make plasma expanders, or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)
XX Sequence 388 AA;
XX Query Match 100.0%; Score 47; DB 1; Length 389;
XX Best Local Similarity 100.0%; Pred. No. 0.42;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKPLEKSH 9
Db 280 EKPLEKSH 288

RESULT 11
AAP90391
ID AAP90391 standard; protein; 390 AA.
XX AC AAP90391;
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX N-terminal human serum albumin.
XX N-terminal portion of human serum albumin; plasma expanders.
XX Homo sapiens; (Human).
XX EP322094-A.
XX 28-JUN-1989.
XX 25-OCT-1988; 88EP-00310000.
XX 30-OCT-1987; 87GB-00025529.
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX WPI; 1989-186464/26.
XX New N-terminal fragments of human serum albumin - esp. useful as blood plasma expanders.
XX Claim 2; Page 9; 20pp; English.
XX N-terminal portion of human serum albumin. Used to make new N-terminal fragments which are used as plasma expanders, or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct PA field)

Qy 1 EKPLEKSH 9
Db 280 EKPLEKSH 288

RESULT 10
AAP90390
ID AAP90390 standard; protein; 389 AA.
XX AC AAP90390;
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX N-terminal human serum albumin.
XX N-terminal portion of human serum albumin; plasma expanders.
XX Homo sapiens; (Human).

```

CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX Sequence 390 AA;
SQ
  Query Match      100.0%; Score 47; DB 1; Length 390;
  Best Local Similarity 100.0%; Pred. No. 0.43; 0; Indels 0; Gaps 0;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9
Db 280 EKPLLEKSH 288

RESULT 12
AAP90392
ID AAP90392 standard; protein; 407 AA.
XX
AC AAP90392;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
DE N-terminal human serum albumin.
XX
KW Human serum albumin; mature protein; new polypeptides; plasma expanders.
XX
OS Homo sapiens; (Human).
XX
FN EP322094-A.
XX
PD 28-JUN-1989.
XX
PF 25-OCT-1988; 88EP-00310000.
XX
PR 30-OCT-1987; 87GB-00025529.
XX
PA (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX
PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PU;
XX
WPI; 1989-186464/26.
XX
New N-terminal fragments of human serum albumin - esp. useful as blood
plasma expanders.
XX
PS Claim 2; Page 9; 20pp; English.
XX
CC N-terminal portion of human serum albumin. Used to make plasma expanders,
or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25
-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
field)
XX
SQ Sequence 407 AA;
  Query Match      100.0%; Score 47; DB 1; Length 407;
  Best Local Similarity 100.0%; Pred. No. 0.45; 0; Indels 0; Gaps 0;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9
Db 280 EKPLLEKSH 288

RESULT 13
AAR14179
ID AAR14179 standard; protein; 463 AA.
XX
AC AAR14179;
XX
DT 19-DEC-1991 (first entry)
XX
DE Human serum albumin lacking N-terminal fragment.

XX HSA; fusion protein; drug.
XX Homo sapiens.
XX JP03201987-A.
XX 03-SEP-1991.
XX 29-DEC-1989; 89JP-00344701.
XX 29-DEC-1989; 89JP-00344701.
XX (TOFU ) TONEN CORP.
XX WPI; 1991-300976/41.
XX
Human serum albumin fragment - where C-terminal of human serum albumin is
lacking and which can be combined with various drugs.
XX Claim 6; Page 1; 23pp; Japanese.
XX
CC This sequence corresponds to amino acids 123 to 585 of mature human serum
albumin. The fragment lacking the N-terminal sequence can form part of a
fusion protein, for example with drugs. (This sequence is taken from the
full-length HSA sequence in EP-330451). See also AAR14178
XX
SQ Sequence 463 AA;
  Query Match      100.0%; Score 47; DB 2; Length 463;
  Best Local Similarity 100.0%; Pred. No. 0.51; 0; Indels 0; Gaps 0;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9
Db 158 EKPLLEKSH 166

RESULT 14
ADD32019
ID ADD32019 standard; protein; 500 AA.
XX
AC ADD32019;
XX
DT 15-JAN-2004 (first entry)
XX
DE Heterologous fusion protein related protein sequence SEQ ID NO:25.
XX
KW heterologous fusion protein;
KW hyperglycosylated granulocyte-colony stimulating factor; G-CSF;
KW human albumin; human albumin analogue; immunoglobulin; Fc;
KW immunostimulant; protein therapy; neutrophil level;
KW insufficient circulating neutrophil level;
KW chronic congenital neutropenia.
XX
OS Synthetic.
XX
FN WO2003076567-A2.
XX
PD 18-SEP-2003.
XX
PF 21-FEB-2003; 2003WO-US003120.
XX
PR 05-MAR-2002; 2002US-0361948P.
XX
PA (ELIL ) LILLY & CO ELI.
XX
PI Beals JM, Kuchibhotla U;
XX
WPI; 2003-902770/82.
XX
New heterologous fusion proteins with granulocyte-colony stimulating
factor activity, useful for increasing neutrophil levels and treating

```

PT patients with low circulating neutrophils, such as after chemotherapy or
 PT in neutropenia.
 XX
 PS Disclosure; SEQ ID NO 25; 126pp; English.
 XX
 CC The present invention describes a heterologous fusion protein (I) comprising a hyperglycosylated granulocyte-colony stimulating factor (G-CSF) analogue fused to a polypeptide having human albumin, human albumin analogues, or fragments of human albumin, or the FC portion of an immunoglobulin, an analogue of the FC portion of an immunoglobulin, or fragments of the FC portion of an immunoglobulin. Also described: (1) a heterologous fusion protein, which is the product of the expression in a host cell of an exogenous DNA sequence encoding (1); (2) an isolated nucleic acid sequence comprising: (a) polynucleotides encoding (1); or (b) a polynucleotide, which comprises any of 15 DNA sequences each comprising 1044 base pairs (see ADD31995 to ADD32010), fused to the DNA encoding a protein (i.e. human albumin, human albumin analogue or fragments of human albumin); (3) increasing neutrophil levels in a mammal comprising the administration of (1); (4) pharmaceutical formulations adapted for the treatment of patients with insufficient neutrophil levels comprising any of (1); (5) a vector comprising the polynucleotide of (2); (6) host cells comprising the vector of (5), or expressing at least one heterologous protein; and (7) producing (1). (1) has immunostimulant activity, and can be used in protein therapy. (1) can be used for increasing neutrophil levels in a mammal. (1) are particularly useful in medicaments for the treatment of patients with insufficient circulating neutrophil levels, or for the manufacture of a medicament for the treatment of patients with insufficient circulating neutrophil levels. The fusion protein is particularly useful for treating conditions treatable by stimulation of circulating neutrophils, such as after chemotherapy regimens or in chronic congenital neutropenia. The present sequence is used in the exemplification of the present invention.

XX Sequence 500 AA;

Query Match 100.0%; Score 47; DB 7; Length 500;
 Best Local Similarity 100.0%; Pred. No. 0.56;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9
 DB 469 EKPLEKSH 477

RESULT 15
 AAU29877
 ID AAU29877 standard; protein; 550 AA.

XX AC AAU29877;

DT 18-DEC-2001 (first entry)

XX DE Novel human secreted protein #368.

XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX FN WO200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US008656.

XX PR 18-APR-2000; 2000US-00552929.

XX PR 26-JAN-2001; 2001US-00770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX

DR WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.

XX PS Claim 20; Page 206; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention

XX Sequence 550 AA;

Query Match 100.0%; Score 47; DB 4; Length 550;
 Best Local Similarity 100.0%; Pred. No. 0.62;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9

DB 245 EKPLEKSH 253

Search completed: April 19, 2004, 11:51:20

Job time : 5.29363 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 3.78947 Seconds
(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_280_288

Perfect score: 47
Sequence: 1 EKPILLEKSH 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	114	12	US-10-296-115-799
2	47	100.0	585	9	Sequence 799, Appl
3	47	100.0	585	10	Sequence 2, Appl
4	47	100.0	585	10	Sequence 445, Appl
5	47	100.0	585	10	Sequence 26, Appl
6	47	100.0	585	10	Sequence 18, Appl
7	47	100.0	585	10	Sequence 18, Appl
8	47	100.0	585	10	Sequence 445, Appl
9	47	100.0	585	11	Sequence 18, Appl
10	47	100.0	585	11	Sequence 18, Appl
11	47	100.0	585	12	Sequence 18, Appl
12	47	100.0	585	12	Sequence 11, Appl
13	47	100.0	585	12	Sequence 31, Appl
14	47	100.0	585	13	Sequence 34, Appl
15	47	100.0	585	14	Sequence 5, Appl

16	47	100.0	585	14	US-10-319-263-1	Sequence 1, Appl
17	47	100.0	585	14	US-10-319-263-2	Sequence 2, Appl
18	47	100.0	585	14	US-10-414-469-1	Sequence 1, Appl
19	47	100.0	585	14	US-10-414-469-2	Sequence 2, Appl
20	47	100.0	585	14	US-10-413-831-1	Sequence 1, Appl
21	47	100.0	585	14	US-10-413-831-2	Sequence 2, Appl
22	47	100.0	585	15	US-10-413-832-1	Sequence 1, Appl
23	47	100.0	585	15	US-10-413-832-2	Sequence 2, Appl
24	47	100.0	585	15	US-10-414-386-1	Sequence 1, Appl
25	47	100.0	585	15	US-10-414-386-2	Sequence 2, Appl
26	47	100.0	585	15	US-10-233-675A-11	Sequence 11, Appl
27	47	100.0	585	15	US-10-462-262-26	Sequence 26, Appl
28	47	100.0	604	10	US-09-984-010-7	Sequence 7, Appl
29	47	100.0	609	10	US-09-919-039-370	Sequence 370, Appl
30	47	100.0	609	12	US-10-609-346-12	Sequence 12, Appl
31	47	100.0	609	13	US-10-153-064-7	Sequence 7, Appl
32	47	100.0	609	14	US-10-153-604B-7	Sequence 7, Appl
33	47	100.0	609	14	US-10-365-623-23	Sequence 23, Appl
34	47	100.0	610	9	US-09-984-186-2	Sequence 2, Appl
35	47	100.0	610	14	US-10-237-667-2	Sequence 2, Appl
36	47	100.0	610	14	US-10-237-708-2	Sequence 2, Appl
37	47	100.0	610	14	US-10-237-866-2	Sequence 2, Appl
38	47	100.0	610	14	US-10-237-871-2	Sequence 2, Appl
39	47	100.0	610	14	US-10-237-624-2	Sequence 2, Appl
40	47	100.0	616	12	US-10-433-108-13	Sequence 13, Appl
41	47	100.0	624	12	US-10-433-108-16	Sequence 16, Appl
42	47	100.0	631	12	US-10-433-108-14	Sequence 14, Appl
43	47	100.0	640	12	US-10-433-108-15	Sequence 15, Appl
44	47	100.0	640	12	US-10-433-108-17	Sequence 17, Appl
45	47	100.0	651	13	US-10-153-064-133	Sequence 133, Appl

ALIGNMENTS

RESULT 1

US-10-296-115-799
; Sequence 799, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: NO. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 799
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(114)
; OTHER INFORMATION: Xaa = any amino acid or other as shown in Table 3
US-10-296-115-799

Query Match 100.0%; Score 47; DB 12; Length 114;

Best Local Similarity 100.0%; Pred. No. 0.28;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPILLEKSH 9

DB 79 EKPILLEKSH 87

RESULT 2

US-09-929-552-2

; Sequence 2, Application US/09929552

; Patent No. US20020123080A1

Query Match 100.0%; Score 47; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 280 EKPLLEKSH 288
|||||

RESULT 10

US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match 100.0%; Score 47; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 280 EKPLLEKSH 288
|||||

RESULT 11

US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match 100.0%; Score 47; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9

Db 280 EKPLLEKSH 288
|||||

RESULT 12

US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanchet, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match 100.0%; Score 47; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 280 EKPLLEKSH 288
|||||

RESULT 13

US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GIP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34

Query Match 100.0%; Score 47; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
Db 280 EKPLLEKSH 288
|||||

RESULT 14

US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 47; DB 13; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9
|||
Db 280 EKPLLEKSH 288

RESULT 15

US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5

Query Match 100.0%; Score 47; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9
|||
Db 280 EKPLLEKSH 288

Search completed: April 19, 2004, 12:54:59
Job time : 3.78947 secs

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 3.47368 Seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_439_447

Perfect score: 51

Sequence: 1 KHPEAKRMP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SFREMEL_25.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	396	4	Q81UK7
2	51	100.0	417	4	Q86YG0
3	39	76.5	125	16	Q8FDR1
4	39	76.5	125	16	Q8FDR1
5	39	76.5	254	2	Q9RLC5
6	39	76.5	352	16	Q91618
7	39	76.5	352	16	Q88QX2
8	39	76.5	352	16	Q88A98
9	38	74.5	55	16	Q92S86
10	38	74.5	98	5	Q9YXG6
11	37	72.5	415	10	Q84VX7
12	37	72.5	765	16	Q89KX8
13	37	72.5	1036	10	Q9C8K6
14	36	70.6	122	9	Q94MS7
15	36	70.6	224	16	Q8A5V9
16	36	70.6	267	16	O67673 aquifex ao

17	36	70.6	321	11	Q8C846
18	36	70.6	375	4	Q8C846 mus musculu
19	36	70.6	412	4	Q9NWF5
20	36	70.6	527	4	Q9NWF5 homo sapien
21	36	70.6	613	4	Q96T01
22	36	70.6	754	4	Q96SM7
23	36	70.6	1485	4	Q9NCE5
24	36	70.6	1486	4	Q9NCE5 homo sapien
25	36	70.6	1644	4	Q9NCE5
26	36	70.6	1685	4	Q9NCE5
27	36	70.6	1740	4	Q9NCE5
28	36	70.6	1758	11	Q8CHN9
29	36	70.6	1902	11	Q925Q1
30	36	70.6	1939	4	Q925Q1
31	36	70.6	1957	4	Q925Q1
32	36	70.6	1999	4	Q9NFD6
33	36	70.6	2165	4	Q9NFD6
34	36	70.6	2285	4	Q9HBJ5
35	35	68.6	68	16	Q8A7F9
36	35	68.6	84	16	Q8A7F9
37	35	68.6	84	16	Q96913
38	35	68.6	125	2	Q8VR61
39	35	68.6	125	2	Q8VR61
40	35	68.6	125	2	Q8VQ03
41	35	68.6	125	2	Q8VQ03
42	35	68.6	125	16	Q8FAS7
43	35	68.6	125	15	Q83JN8
44	35	68.6	153	13	Q8QSK3
45	35	68.6	182	5	Q8MZY6

ALIGNMENTS

RESULT 1

Q81UK7 PRELIMINARY; PRT; 396 AA.

AC Q81UK7; TREMBLrel. 23, Created

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 25, Last annotation update)

DE Similar to serum albumin precursor.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1] SEQUENCE FROM N.A.

RP TISSUE=Liver;

RA Strausberg R.;

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC035969; AAH35969.1; -

DR GO; GO:0005386; C:extracellular space; IEA.

DR GO; GO:0005386; F:carrier activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; transport_prot; 2.

DR PRINTS; PR00802; SERUMALBUMIN.

DR SMART; SM0103; ALBUMIN; 2.

DR PROSITE; PS00212; ALBUMIN; 2.

SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 51; DB 4; Length 396;

Best Local Similarity 100.0%; Pred. No. 0.21;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9

Db 250 KHPEAKRMP 258

RESULT 2

Q86YG0

ID Q86YGO PRELIMINARY; PRT; 417 AA.
 AC Q86YGO;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to alpha-fetoprotein.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (DSC-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC041789; AA41789.1;
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; F:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport prot.; 2.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR PRODOM; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 2.
 DR PROSITE; PS00212; ALBUMIN; 2.
 SQ SEQUENCE 417 AA; 47360 MW; 16E764833BEF4E8D CRC64;

Query Match 100.0%; Score 51; DB 4; Length 417;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KHPEAKR 9
 Db 271 KHPEAKRMP 279

RESULT 3
 Q86YGO PRELIMINARY; PRT; 125 AA.
 ID Q86YGO
 AC Q86YGO;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein yeev.
 GN C3677.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL; AE016766; AAN82125.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 125 AA; 14014 MW; 26AB67E398121996 CRC64;

Query Match 76.5%; Score 39; DB 16; Length 125;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KHPEAKR 7
 Db 119 KHPEAKR 125

RESULT 4
 Q86YGO PRELIMINARY; PRT; 125 AA.
 ID Q86YGO
 AC Q86YGO;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein yeev.
 GN C4576.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL; AE016769; AAN83010.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 125 AA; 13964 MW; 473BED42C8398644 CRC64;

Query Match 76.5%; Score 39; DB 16; Length 125;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KHPEAKR 7
 Db 119 KHPEAKR 125

RESULT 5
 Q86YGO PRELIMINARY; PRT; 254 AA.
 ID Q86YGO
 AC Q86YGO;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Biotin synthetase (EC 2.8.1.6) (Fragment)
 GN B10B.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JM300;
 RA Graupner S., Wackernagel W.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ249742; CAB56476.1;
 DR GO; GO:0004076; F:biotin synthase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0009102; P:biotin biosynthesis; IEA.
 DR InterPro; IPR002684; Biotin_synth.
 DR InterPro; IPR006638; Elp3.
 DR InterPro; IPR007197; Radical SAM.
 DR Pfam; PF04055; Radical SAM; 1.
 DR SMART; SM00729; Elp3; 1.
 DR TIGRFAMs; TIGR00433; bioB; 1.
 KW Transferase.
 FT NON TER 254 254
 SQ SEQUENCE 254 AA; 27771 MW; F512DF0FEFC12F67 CRC64;

Query Match 76.5%; Score 39; DB 2; Length 254;
 Best Local Similarity 77.8%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 KHPEAKRMP 9
Db 109 KHPSAKDMP 117

RESULT 6
Q91618 ID Q91618 PRELIMINARY; PRT; 352 AA.
AC Q91618;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Biotin synthetase.
GN BIOC OR PA0500.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Raizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE004487; AAG03889.1; -.
DR PIR; F83582; F83582.
DR GO; GO:0004076; F:biotin synthase activity; IEA.
DR GO; GO:0009102; F:biotin biosynthesis; IEA.
DR InterPro; IPR002684; Biotin_synth.
DR InterPro; IPR006638; BIP3.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; I.
DR SMART; SM00729; Elp3; I.
DR TIGRFAMs; TIGR00433; bioB; I.
KW Complete proteome.
SQ SEQUENCE 352 AA; 39113 MW; E46491B67DDF369A CRC64;

Query Match 76.5%; Score 39; DB 16; Length 352;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 109 KHPSAKDMP 117

RESULT 7
Q88QX2 ID Q88QX2 PRELIMINARY; PRT; 352 AA.
AC Q88QX2;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Biotin synthetase.
GN BIOC OR PF0362.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,

Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016775; AAN65993.1; -.
DR TIGR; PP0362; -.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; I.
KW Complete proteome.
SQ SEQUENCE 352 AA; 38950 MW; 7EAC944861D3553F CRC64;

Query Match 76.5%; Score 39; DB 16; Length 352;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 109 KHPSAKDMP 117

RESULT 8
Q88A98 ID Q88A98 PRELIMINARY; PRT; 352 AA.
AC Q88A98;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Biotin synthetase.
GN BIOC OR PSPT00494.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Daviden T.,
RA White O., Fraser C., Collmer A.;
RT "Complete sequence of Pseudomonas syringae."
RT Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AE016857; AAC34038.1; -.
DR TIGR; PSPT00494; -.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; I.
KW Complete proteome.
SQ SEQUENCE 352 AA; 38610 MW; B2AB050D00F03F50 CRC64;

Query Match 76.5%; Score 39; DB 16; Length 352;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
Db 109 KHPSAKDMP 117

RESULT 9
Q9S2S6 ID Q9S2S6 PRELIMINARY; PRT; 55 AA.
AC Q9S2S6;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Putative small hydrophilic protein.
GN SCO2063 OR SC4G6.32.
OS Streptomyces coelicolor.

```

```
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=A3(2) / X145;
MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Cronin A., Fraser A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Crispin A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL039111; CAB51455.1; -
DR FIR; T35092; T35092.
KW Complete proteome.
SQ SEQUENCE 55 AA; 6247 MW; 2F34EDC4E906354F CRC64;

Query Match 74.5%; Score 38; DB 16; Length 55;
Best Local Similarity 66.7%; Pred. No. 8.2;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9
DB 11 KHPQAKRSP 19

RESULT 10
Q9GVG6 PRELIMINARY; PRT; 98 AA.
AC Q9GVG6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN W01C8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nhan M.;
RT "The sequence of C. elegans cosmid W01C8.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U41508; AAG00023.1; -
DR FIR; T26046; T26046.
DR WormPep; W01C8.1; CE05029.
KW Hypothetical protein.
SQ SEQUENCE 98 AA; 11596 MW; D65B2FFBF9FB859 CRC64;

Query Match 74.5%; Score 38; DB 5; Length 98;

Best Local Similarity 75.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 2 HPEAKRMP 9
DB 13 HPEEKRLP 20

RESULT 11
Q84VX7 PRELIMINARY; PRT; 415 AA.
AC Q84VX7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Atg5l540.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=Col-0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN W01C8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Nhan M.;
RT "The sequence of C. elegans cosmid W01C8.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct submission.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U41508; AAG00023.1; -
DR FIR; T26046; T26046.
DR WormPep; W01C8.1; CE05029.
KW Hypothetical protein.
SQ SEQUENCE 98 AA; 11596 MW; D65B2FFBF9FB859 CRC64;

Query Match 74.5%; Score 38; DB 5; Length 98;

Best Local Similarity 75.0%; Pred. No. 95;
Matches 6; Conservative 1; Mismatches 0; Gaps 0;

QY 1 KHPKAKRMP 8
DB 338 KHPETKKM 345

RESULT 12
Q89KX8 PRELIMINARY; PRT; 765 AA.
AC Q89KX8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BIR4770 protein.
GN BIR4770.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN=USDA 110;
MEDLINE=22484999; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasekawa S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Teurloka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
```

```

RL DNA Res. 9:189-197(2002).
DR EMBL; APO05952; BAC0035.1; -.
DR InterPro; IPR008258; SLT dom.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF01464; SLT; 1.
KW Complete proteome.
SQ SEQUENCE 765 AA; 84473 MW; 8F26CB2DEE360326 CRC64;

Query Match 72.5%; Score 37; DB 16; Length 765;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9
DB 94 KHAFAKRP 102

RESULT 13
Q9C8K6 PRELIMINARY; PRT; 1036 AA.
AC Q9C8K6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F5D21.24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia.
RA MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altarfi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egtu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC024261; AAG52629.1; -.
DR PIR; H96553; H96553.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR006652; Kelch rep.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF01344; Kelch; 4.
DR PROSITE; PS00334; MYB_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 1036 AA; 118234 MW; C73FE8E7A802472C CRC64;

Query Match 72.5%; Score 37; DB 10; Length 1036;
Best Local Similarity 75.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 8
DB 959 KHPETKRM 966

```

```

RESULT 14
Q94MS7 PRELIMINARY; PRT; 122 AA.
AC Q94MS7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE P42.
OS Bacteriophage Mx8.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC P22-like viruses.
OX NCBI_TaxID=49964;
RN [1]
RP SEQUENCE FROM N.A.
RA Youderian P., Walthers D., Salmi D., Magrini V., Hartzell P.L.;
RT "Genome organization of temperate Myxococcus phage Mx8.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF396866; AAK94377.1; -.
SQ SEQUENCE 122 AA; 12811 MW; 6F3E8978D3ADB17A CRC64;

Query Match 70.6%; Score 36; DB 9; Length 122;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HPEAKRM 8
DB 53 HPEAKRL 59

RESULT 15
Q8ASV9 PRELIMINARY; PRT; 224 AA.
AC Q8ASV9
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative beta-phosphoglucomutase.
GN BZ1217.
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VP1-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12563928;
RA Xu J., Bjursell M.K., Hamrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AS016934; AAQ77234.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006402; HAD-SF-IA-v3.
DR InterPro; IPR005833; Hlgase/hydrolase.
DR InterPro; IPR005834; Hydrolase.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00413; HADHLOGNASE.
DR TIGRfams; TIGR01509; HAD-SF-IA-v3; 1.
KW Complete proteome.
SQ SEQUENCE 224 AA; 24939 MW; DF4ECF191D09538 CRC64;

Query Match 70.6%; Score 36; DB 16; Length 224;
Best Local Similarity 75.0%; Pred. No. 79;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HPEAKRMP 9
DB 86 YPEAKRMP 93

```

Search completed: April 19, 2004, 12:00:13

Job time : 5.47368 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 5.29363 Seconds

(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_439_447

Perfect score: 51

Sequence: 1 KHPKAKMP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	106	4 AAO04431	Aao04431 Human pol
2	51	100.0	119	4 AAO02645	Aao02645 Human pol
3	51	100.0	133	4 AAO03606	Aao03606 Human pol
4	51	100.0	151	4 AAU33085	Aau33085 Novel hum
5	51	100.0	187	4 AAU27873	Aau27873 Human con
6	51	100.0	228	3 AAY83949	Aay83949 Yeast cod
7	51	100.0	245	4 AAU33081	Aau33081 Novel hum
8	51	100.0	386	4 AAU33074	Aau33074 Novel hum
9	51	100.0	401	4 AAU23876	Aau23876 Novel hum
10	51	100.0	507	4 AAU33286	Aau33286 Novel hum
11	51	100.0	507	4 AAU32994	Aau32994 Novel hum
12	51	100.0	550	4 AAU29877	Aau29877 Novel hum
13	51	100.0	584	6 ABG72381	Abg72381 Mature hu
14	51	100.0	585	1 AAP93344	Aap93344 Sequence
15	51	100.0	585	1 AAP93344	Aap93344 Mature hu
16	51	100.0	585	2 AAP90388	Aap90388 Human ser
17	51	100.0	585	2 AAR03318	Aar03318 Human ser
18	51	100.0	585	2 AAR08457	Aar08457 Human ser
19	51	100.0	585	2 AAR26207	Aar26207 Human ser
20	51	100.0	585	2 AAR26362	Aar26362 Synthetic
21	51	100.0	585	2 AAR20029	Aar20029 Human ser
22	51	100.0	585	2 AAR80301	Aar80301 Human ser
23	51	100.0	585	2 AAO20111	Aao20111 HSA prote
24	51	100.0	585	2 AAWS5841	Aaw5841 Mature pr
25	51	100.0	585	3 AAY84873	Aay84873 Amino aci
26	51	100.0	585	3 AAY83946	Aay83946 Yeast cod

ALIGNMENTS

RESULT 1

AAO04431

ID AAO04431 standard; protein; 106 AA.

XX AC AAO04431;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 18323.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX PS WPI; 2001-514838/56.

XX PS N-PSDB; AAI84362.

XX PS Claim 20; SEQ ID NO 18323; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and

XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC isolated nucleic acids and polypeptides, useful for preventing diagnosing

XX CC and treating e.g. leukemia, inflammation and immune disorders.

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 106 AA;

Query Match 100.0%; Score 51; DB 4; Length 106;
 Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
 |||||
 Db 65 KHPEAKRMP 73

RESULT 2
 AAO02645
 ID AAO02645 standard; protein; 119 AA.

AC AAO02645;

XX 06-NOV-2001 (first entry)

DT Human polypeptide SEQ ID NO 16537.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AA182576.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 16537; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 119 AA;

Query Match 100.0%; Score 51; DB 4; Length 119;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
 |||||
 Db 46 KHPEAKRMP 54

RESULT 3

AAO03606

ID AAO03606 standard; protein; 133 AA.

XX AAO03606;

XX 06-NOV-2001 (first entry)

DT Human polypeptide SEQ ID NO 17498.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AA183537.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 17498; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 133 AA;

Query Match 100.0%; Score 51; DB 4; Length 133;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
 |||||
 Db 78 KHPEAKRMP 86

RESULT 4

AAU33085

ID AAU33085 standard; protein; 151 AA.

XX AAU33085;

XX 18-DEC-2001 (first entry)
XX Novel human secreted protein #3576.
XX Human; vaccination; Gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX Homo sapiens.
XX WO200179449-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US008656.
XX 18-APR-2000; 2000US-00552929.
XX 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX Claim 20; Page 705; 765pp; English.
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising the
XX nucleic acids encoding the polypeptides and cells genetically engineered
XX to express them are also useful for producing the proteins. The proteins
XX are useful in genetic vaccination, testing and therapy, and can be used
XX as nutritional supplements. They may be used to increase stem cell
XX proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX and/or nerve tissue growth or regeneration; immune suppression and/or
XX stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX AAU29510-AAU3304 represent the amino acid sequences of novel human
XX secreted proteins of the invention
XX Sequence 151 AA;
XX Query Match 100.0%; Score 51; DB 4; Length 151;
XX Best Local Similarity 100.0%; Pred. NO. 0.18;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 KHPKAKRMP 9
XX Db 12 KHPKAKRMP 20
XX
XX RESULT 5
XX AAU27873
XX ID AAU27873 standard; protein; 187 AA.
XX AC AAU27873;
XX 18-DEC-2001 (first entry)
XX Human contig polypeptide sequence #26.
XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
XX mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
XX cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;

XX nervous system disorder; inflammatory disorder; cell differentiation;
XX angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
XX genetic disorder; bone regeneration; tendon; ligament; tissue repair;
XX cytostatic; antirheumatic; antiarthritic; vulnery; antiinflammatory;
XX antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
XX neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
XX immunostimulant; analgesic; gene therapy.
XX Homo sapiens.
XX Synthetic.
XX WO200164834-A2.
XX 07-SEP-2001.
XX 26-FEB-2001; 2001WO-US004926.
XX 28-FEB-2000; 2000US-00515126.
XX 18-MAY-2000; 2000US-00577409.
XX 17-JUN-2000; 2000US-00597707.
XX 14-JUL-2000; 2000US-00616807.
XX 19-SEP-2000; 2000US-00664641.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
XX Drmanac R;
XX WPI; 2001-589862/66.
XX N-PSDB; AAS44773.
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
XX prepared from various human tissues, for diagnosis, treatment of cancer,
XX neurological, inflammatory disorders and for use in arrays for detection.
XX Claim 10; Page 127; 153pp; English.
XX Sequences AAU27676-AAU28019 represent full-length polypeptides and contig
XX polypeptides of the invention. The proteins and their associated DNA
XX sequences are useful for the treatment, diagnosis and prevention of
XX various types of disorder in a mammalian subject such as a human, dog,
XX monkey, mouse, hamster or rat. The disorders include cancers such as
XX leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as
XX multiple sclerosis, connective tissue disease, rheumatoid arthritis,
XX diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
XX disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
XX chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
XX Wernicke disease, inflammatory disorders such as nephritis, Crohn's
XX disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
XX bowel disease. The sequences exhibit activity relating to angiogenesis,
XX cell proliferation, cell differentiation, stem cell growth factor,
XX activin or inhibin. Therefore, they can be used to manipulate stem cells
XX in culture to give rise to neuroepithelial cells that can be used to
XX augment or replace cells damaged by illness, accidental damage or genetic
XX disorders. The sequences may also be used for regeneration of bone,
XX cartilage, tendons and ligaments and in tissue repair and burn healing.
XX Note: Some sequences for this patent did not form part of the printed
XX at ftp.wipo.int/pub/published_pct_sequences

Sequence 187 AA;

Query Match 100.0%; Score 51; DB 4; Length 187;

Best Local Similarity 100.0%; Pred. NO. 0.22; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPKAKRMP 9

Db 75 KHPKAKRMP 84

RESULT 6

AAV83949
ID AAY83949 standard; protein; 228 AA.
XX
AC AAY83949;
XX
DT 28-JUL-2000 (first entry)
XX
DE Yeast codon-biased recombinant HSA protein fragment HSA-III.
XX
KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
KW overlapping oligonucleotide; expression vector.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CN1239103-A.
XX
PD 22-DEC-1999.
XX
PF 17-JUN-1998; 98CN-00102506.
XX
PR 17-JUN-1998; 98CN-00102506.
XX
PA (HALJ-) HALJI BIOENGINEERING CO LTD.
XX
PI Li S, Lu D;
XX
DR WPI; 2000-351198/31.
DR N-PSDB; AAA10094.
XX
PT Process for preparing recombinant human serum albumin comprising yeast
PT biased sex codons - uses a recombinant DNA technique.
XX
PS Example 1; Fig 7; 44pp; Chinese.
XX
CC The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AA10091) was generated as
CC three synthetic fragments (AA10092-AA10094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the sequence of the HSA fragment HSA-III encoded by the human gene with a
CC yeast codon bias. The invention also covers a recombinant expression
CC vector, yeast host cells carrying the recombinant expression vector and
CC the process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
SQ Sequence 228 AA;
Query Match 100.0%; Score 51; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KXPEAKRMP 9
DB 82 KXPEAKRMP 90
RESULT 7
AAU33081
ID AAU33081 standard; protein; 245 AA.
XX
AC AAU33081;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3572.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.

XX WO200179449-A2.
PN
XX 25-OCT-2001.
PD
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-611725/70.
DR
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
PT
XX
XX Claim 20; Page 705; 765pp; English.
PS
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 245 AA;
Query Match 100.0%; Score 51; DB 4; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KXPEAKRMP 9
DB 167 KXPEAKRMP 175
RESULT 8
AAU33074
ID AAU33074 standard; protein; 386 AA.
XX
AC AAU33074;
XX
XX 18-DEC-2001 (first entry)
DT
XX
XX Novel human secreted protein #3565.
DE
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
OS
XX WO200179449-A2.
PN
XX 25-OCT-2001.
PD
XX 16-APR-2001; 2001WO-US008656.
PF
XX 18-APR-2000; 2000US-00552929.
PR

```

PR 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 704; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 386 AA;
    Query Match          100.0%; Score 51; DB 4; Length 386;
    Best Local Similarity 100.0%; Pred. No. 0.49;
    Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
   |||||
DB 159 KHPEAKRMP 167

RESULT 9
AAU29876
ID AAU29876 standard; protein; 401 AA.
XX
AC AAU29876;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #367.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 386 AA;
    Query Match          100.0%; Score 51; DB 4; Length 386;
    Best Local Similarity 100.0%; Pred. No. 0.49;
    Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KHPEAKRMP 9
   |||||
DB 159 KHPEAKRMP 167

RESULT 10
AAU33286
ID AAU33286 standard; protein; 507 AA.
XX
AC AAU33286;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3777.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX
CC Nucleic acids encoding a range of human polypeptides, useful in genetic
CC vaccination, testing and therapy.
XX
PS Claim 20; Page 754; 765pp; English.
XX
CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated

```

CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 507 AA;

Query Match 100.0%; Score 51; DB 4; Length 507;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
|||||
Db 215 KHPEAKRMP 223

RESULT 11
AAU32994
ID AAU32994 standard; protein; 507 AA.

AC AAU32994;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #3485.

KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US008656.

PR 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.

PS Claim 20; Page 696; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used

CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 507 AA;

Query Match 100.0%; Score 51; DB 4; Length 507;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
|||||
Db 215 KHPEAKRMP 223

RESULT 12
AAU29877

ID AAU29877 standard; protein; 550 AA.

AC AAU29877;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #368.

KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US008656.

PR 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.

PS Claim 20; Page 206; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU33304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 550 AA;

```

Query Match      100.0%; Score 51; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
DB      404 KHPEAKRMP 412

RESULT 13
ABG72381
ID ABG72381 standard; protein; 584 AA.
XX
AC ABG72381;
XX
DT 10-FEB-2003 (first entry)
DE
XX
DE Mature human serum albumin #2.
XX
KW Human; serum albumin; HSA; cancer; cytostatic; breast cancer;
KW prostate cancer; anti-proliferative.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 241..242
FT /note= "GTCCACACG"
XX
PN US2002123080-A1.
XX
PD 05-SEP-2002.
XX
PF 14-AUG-2001; 2001US-00929552.
XX
PR 19-DEC-1996; 96US-00769746.
XX
PA (TUFT ) UNIV TUFTS.
XX
PI Sonnenschein C, Soto AM;
XX
DR WPI: 2003-066789/06.
DR N-PSDB; ABX13582.
XX
FT Testing human cancer cells, especially breast and prostate cancer cells,
FT by contacting cancer cells obtained from biopsy of a patient ex vivo with
FT antibody specific to human albumin receptor.
XX
PS Disclosure; Fig 2; 17pp; English.
XX
CC The invention relates to testing human cancer cells, comprising obtaining
CC cancer cells from the patient and contacting the cell ex vivo with an
CC antibody to the receptor for human albumin. The method is useful for
CC testing human cancer cells in particular breast and prostate cancer
CC cells. The present sequence is mature human serum albumin, HSA. The anti-
CC proliferative effect of HSA was assayed in an experiment included in the
CC specification
XX
SQ      Sequence 584 AA;

Query Match      100.0%; Score 51; DB 6; Length 584;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
DB      438 KHPEAKRMP 446

RESULT 14
AAP93344
ID AAP93344 standard; protein; 585 AA.
XX
AC AAP93344;
XX
DT 25-MAR-2003 (revised)
DT 23-JUN-1990 (first entry)
XX
DE Sequence of mature human serum albumin (HSA) as encoded by artificial
DE gene.
XX
KW Mature human serum albumin; artificial gene; oligonucleotide block;
KW hypobolaemia; shock; hypoalbuminaemia.
XX
OS Homo sapiens.
XX
PN EP308381-A.
XX
PD 22-MAR-1989.
XX
PF 13-SEP-1988; 88EP-00850299.
XX
PR 14-SEP-1987; 87SE-00003539.
XX
PA (SKAN-) SKANDIGEN AB.
PA (MAGY ) MTA SZEGEDI BIOLOG KOEZFONTI.
PA (VEPE-) VEPEX CONTRACTOR LTD.
XX
PI Aberg B, Simoncsits A, Kalan M, Csperpan I, Bajszar G;
XX
DR WPI: 1989-087749/12.
DR N-PSDB; AAN90997.
XX
FT Artificial gene coding for authentic human serum albumin - constructed on
FT the basis of codons most frequently used by chosen non-human host.
XX
PS Disclosure; pp. 11-16; 121pp; English.
XX
CC The synthetic gene was constructed by designing a nucleotide sequence in
CC which the codons which are most frequently used by the chosen non-human
CC host were selected. In this case, it is yeast cells (LL20; Leu2-3; 112,
CC His 3-11, 15). The synthetic HSA gene was assembled from 24
CC oligonucleotide blocks. HSA is used in therapy for the treatment of
CC hypovolaemia, shock and hypoalbuminaemia. It is also used as an additive
CC in perfusion liq. for extracorporeal circulation and as an experimental
CC antigen. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR
CC -2003 to correct PI field.)
XX
SQ      Sequence 585 AA;

Query Match      100.0%; Score 51; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
DB      439 KHPEAKRMP 447

RESULT 15
AAP90388
ID AAP90388 standard; protein; 585 AA.
XX
AC AAP90388;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Mature human serum albumin polypeptide.
XX
KW Human serum albumin; mature protein; new polypeptides; plasma expanders.
XX
OS Homo sapiens; (Human).
XX
PN EP322094-A.
```

XX 28-JUN-1989.
PD
XX
XX 25-OCT-1988; 88EP-00310000.
PF
XX
XX 30-OCT-1987; 87GB-00025529.
PR
XX
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
PA
XX
XX Balance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
PI
XX
XX WPI; 1989-186464/26.
DR
XX N-PSDB; AAN90128.
XX
XX New N-terminal fragments of human serum albumin - esp. useful as blood
PI plasma expanders.
XX
XX Disclosure; Fig 2; 20pp; English.
PS
XX Mature protein of human serum albumin (see corresp. AAN90128). Used to
CC make new N-terminal fragments which are used as plasma expanders, or as
CC substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 51; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447
|||||

Search completed: April 19, 2004, 11:51:21
Job time : 6.29363 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 3.78947 Seconds
(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_439_447

Perfect score: 51
Sequence: 1 KHEAKRMP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	585	9	US-09-929-552-2
2	51	100.0	585	10	US-09-932-613-445
3	51	100.0	585	10	US-09-984-010-26
4	51	100.0	585	10	US-09-833-041-18
5	51	100.0	585	10	US-09-833-117-18
6	51	100.0	585	10	US-09-932-322-445
7	51	100.0	585	10	US-09-832-501-18
8	51	100.0	585	11	US-09-833-118-18
9	51	100.0	585	11	US-09-833-245-18
10	51	100.0	585	12	US-10-424-999-11
11	51	100.0	585	12	US-10-425-000-31
12	51	100.0	585	12	US-10-433-108-34
13	51	100.0	585	13	US-10-153-064-5
14	51	100.0	585	14	US-10-153-604A-5
15	51	100.0	585	14	US-10-319-263-1

RESULT 1

US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids

ALIGNMENTS

16	51	100.0	585	14	US-10-319-263-2	Sequence 2, Appli
17	51	100.0	585	14	US-10-414-469-1	Sequence 1, Appli
18	51	100.0	585	14	US-10-414-469-2	Sequence 2, Appli
19	51	100.0	585	14	US-10-413-831-1	Sequence 1, Appli
20	51	100.0	585	14	US-10-413-831-2	Sequence 2, Appli
21	51	100.0	585	15	US-10-413-832-1	Sequence 1, Appli
22	51	100.0	585	15	US-10-413-832-2	Sequence 2, Appli
23	51	100.0	585	15	US-10-414-386-1	Sequence 1, Appli
24	51	100.0	585	15	US-10-414-386-2	Sequence 2, Appli
25	51	100.0	585	15	US-10-233-675A-11	Sequence 11, Appli
26	51	100.0	585	15	US-10-462-262-26	Sequence 26, Appli
27	51	100.0	604	10	US-09-984-010-7	Sequence 7, Appli
28	51	100.0	609	10	US-09-919-039-370	Sequence 370, App
29	51	100.0	609	12	US-10-609-346-12	Sequence 12, Appli
30	51	100.0	609	13	US-10-153-064-7	Sequence 7, Appli
31	51	100.0	609	14	US-10-153-604A-7	Sequence 7, Appli
32	51	100.0	609	14	US-10-365-623-23	Sequence 23, Appli
33	51	100.0	610	9	US-09-984-186-2	Sequence 2, Appli
34	51	100.0	610	14	US-10-237-667-2	Sequence 2, Appli
35	51	100.0	610	14	US-10-237-708-2	Sequence 2, Appli
36	51	100.0	610	14	US-10-237-866-2	Sequence 2, Appli
37	51	100.0	610	14	US-10-237-871-2	Sequence 2, Appli
38	51	100.0	610	14	US-10-237-824-2	Sequence 2, Appli
39	51	100.0	616	12	US-10-433-108-13	Sequence 13, Appli
40	51	100.0	624	12	US-10-433-108-16	Sequence 16, Appli
41	51	100.0	631	12	US-10-433-108-14	Sequence 14, Appli
42	51	100.0	640	12	US-10-433-108-15	Sequence 15, Appli
43	51	100.0	640	12	US-10-433-108-17	Sequence 17, Appli
44	51	100.0	651	13	US-10-153-064-133	Sequence 133, App
45	51	100.0	651	14	US-10-153-604A-133	Sequence 133, App


```
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match      100.0%; Score 51; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 6
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BlyS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932.322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-322-445

Query Match      100.0%; Score 51; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 7
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832.501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 51; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 8
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833.118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 51; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 9
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833.245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18
```

```
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 51; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 8
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833.118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 51; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447

RESULT 9
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833.245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18
```

Query Match 100.0%; Score 51; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447

RESULT 10
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040082810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match 100.0%; Score 51; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447

RESULT 11
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040082777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match 100.0%; Score 51; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9

Db 439 KHPEAKRMP 447

RESULT 12
US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34

Query Match 100.0%; Score 51; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447

RESULT 13
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 51; DB 13; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KHPEAKRMP 9
Db 439 KHPEAKRMP 447

RESULT 14
US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212

```
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version.3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5

Query Match      100.0%; Score 51; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. NO. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447
```

```
RESULT 15
US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISCO07
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-319-263-1
```

```
Query Match      100.0%; Score 51; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. NO. 0.98;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KHPEAKRMP 9
Db      439 KHPEAKRMP 447
```

Search completed: April 19, 2004, 12:54:59
Job time : 3.78947 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.3795 Seconds
(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_280_288

Perfect score: 47

Sequence: 1 EKPLLEKSH 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/prodata/2/aaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/aaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/aaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/aaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/aaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	100.0	585	1 US-08-153-799-14	Sequence 14, Appli
2	47	100.0	585	1 US-08-448-196A-3	Sequence 3, Appli
3	47	100.0	585	2 US-08-984-176-1	Sequence 1, Appli
4	47	100.0	585	2 US-08-702-572-2	Sequence 2, Appli
5	47	100.0	585	3 US-08-769-746-2	Sequence 5, Appli
6	47	100.0	585	4 US-10-153-064-5	Sequence 3, Appli
7	47	100.0	609	1 US-08-222-619-3	Sequence 4, Appli
8	47	100.0	609	1 US-08-433-037-4	Sequence 2, Appli
9	47	100.0	609	4 US-08-897-956A-2	Sequence 7, Appli
10	47	100.0	609	4 US-10-153-064-7	Sequence 977, App
11	47	100.0	609	4 US-09-976-594-977	Sequence 3, Appli
12	47	100.0	609	5 PCT-US95-04075-3	Sequence 2, Appli
13	47	100.0	610	2 US-08-797-689-2	Sequence 2, Appli
14	47	100.0	610	4 US-09-984-186-2	Sequence 2, Appli
15	47	100.0	651	4 US-10-153-064-133	Sequence 133, App
16	47	100.0	652	4 US-10-153-064-96	Sequence 96, Appl
17	47	100.0	652	4 US-10-153-064-99	Sequence 99, Appl
18	47	100.0	652	4 US-10-153-064-105	Sequence 105, App
19	47	100.0	652	4 US-10-153-064-132	Sequence 132, App
20	47	100.0	656	4 US-10-153-064-131	Sequence 131, App
21	47	100.0	656	4 US-10-153-064-130	Sequence 130, App
22	47	100.0	660	4 US-10-153-064-90	Sequence 90, Appl
23	47	100.0	660	4 US-10-153-064-93	Sequence 93, Appl
24	47	100.0	668	4 US-10-153-064-102	Sequence 102, App
25	47	100.0	676	4 US-10-153-064-95	Sequence 95, Appl
26	47	100.0	676	4 US-10-153-064-98	Sequence 98, Appl
27	47	100.0	676	4 US-10-153-064-104	Sequence 104, App

Sequence 127, App
Sequence 129, App
Sequence 125, App
Sequence 123, App
Sequence 92, Appl
Sequence 101, App
Sequence 2, Appli
Sequence 4, Appli
Sequence 16, Appli
Sequence 16, Appli
Sequence 3, Appli
Sequence 89, Appli
Sequence 1, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 5, Appli
Sequence 6, Appli
Sequence 5132, Ap

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:

NAME/KEY: Region
LOCATION: 369..419
OTHER INFORMATION: /note= "Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..585
OTHER INFORMATION: /note= "Amino acid sequence of
OTHER INFORMATION: natural HSA"
US-08-153-799-14

Query Match 100.0%; Score 47; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.33; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPJLEKSH 9
| | | | | | | | | |
DB 280 EKPJLEKSH 288

RESULT 2

US-08-448-196A-3
Sequence 3, Application US/08448196A
Patent No. 5780594

GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C.
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
TITLE OF INVENTION: RELATED PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NASA
STREET: MARSHALL SPACE FLIGHT CENTER
CITY: HUNTSVILLE
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal

US-08-448-196A-3
Query Match 100.0%; Score 47; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.33; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPJLEKSH 9
| | | | | | | | | |
DB 280 EKPJLEKSH 288

RESULT 3

US-08-984-176-1
Sequence 1, Application US/08984176
Patent No. 5948609

GENERAL INFORMATION:
APPLICANT: CARTER, DANIEL C
APPLICANT: HO, JOSEPH X
TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
FILE REFERENCE: 08/984,176
CURRENT APPLICATION NUMBER: US/08/984,176
CURRENT FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 585
TYPE: PRT
ORGANISM: Homo sapiens
US-08-984-176-1

Query Match 100.0%; Score 47; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.33; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPJLEKSH 9
| | | | | | | | | |
DB 280 EKPJLEKSH 288

RESULT 4

US-08-702-572-2
Sequence 2, Application US/08702572
Patent No. 5965386

GENERAL INFORMATION:
APPLICANT: Kerry-Williams, Sean M
APPLICANT: Gilbert, Sarah C
TITLE OF INVENTION: Yeast Strains and Modified Albumins
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Centeon L.L.C.
STREET: 1020 First Avenue
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-1310

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,572
FILING DATE: 11-NOV-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/23857
FILING DATE: 1-MAR-1995
APPLICATION NUMBER: GB 9404270.2
FILING DATE: 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Naomi Biswas
REGISTRATION NUMBER: 38,384
REFERENCE/DOCKET NUMBER: CS0114 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 47; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9
Db 280 EKPLEKSH 288

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Scto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 337-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match 100.0%; Score 47; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9
Db 280 EKPLEKSH 288

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: P2356

; CURRENT APPLICATION NUMBER: US/10/153.064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 47; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9
Db 280 EKPLEKSH 288

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match 100.0%; Score 47; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9
Db 304 EKPLEKSH 312

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan

; SEQUENCE CHARACTERISTICS:
; LENGTH: 595 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 47; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9
Db 280 EKPLEKSH 288

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Scto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 337-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match 100.0%; Score 47; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9
Db 280 EKPLEKSH 288

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: P2356

; CURRENT APPLICATION NUMBER: US/10/153.064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 47; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9
Db 280 EKPLEKSH 288

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match 100.0%; Score 47; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLEKSH 9
Db 304 EKPLEKSH 312

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan

APPLICANT: Barr, Kathryn A.
APPLICANT: Brierley, Russell A.
APPLICANT: Thill, Gregory P.
APPLICANT: Techopp, Jurgis F.
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
TITLE OF INVENTION: PICHIA PASTORIS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,037
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,345
REFERENCE/DOCKET NUMBER: 91082
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 100.0%; Score 47; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.34; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 EKPLLEKSH 9
DB 304 EKPLLEKSH 312

RESULT 9
US-08-897-956A-2
Sequence 2, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/CPA
CURRENT APPLICATION NUMBER: US/08/897,956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 609
TYPE: PRT
ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
DB 304 EKPLLEKSH 312
RESULT 10
US-10-153-064-7
Sequence 7, Application US/10153064
Patent No. 6663485
GENERAL INFORMATION:
APPLICANT: Bell et al.
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
FILE REFERENCE: PFS56
CURRENT APPLICATION NUMBER: US/10/153,064
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: 60/293,212
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 137
SOFTWARE: Patent in version 3.1
SEQ ID NO 7
LENGTH: 609
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match 100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.34; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 EKPLLEKSH 9
DB 304 EKPLLEKSH 312

RESULT 11
US-09-976-594-977
Sequence 977, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 977
LENGTH: 609
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match 100.0%; Score 47; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.34; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 EKPLLEKSH 9
DB 304 EKPLLEKSH 312

RESULT 12
PCI-US95-04075-3
Sequence 3, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```



```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match      100.0%; Score 47; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EKPLLEKSH 9
Db      304 EKPLLEKSH 312

RESULT 15
US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match      100.0%; Score 47; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EKPLLEKSH 9
Db      346 EKPLLEKSH 354

```

Search completed: April 19, 2004, 12:05:20
Job time : 1.3795 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 0.953186 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_362_368

Perfect score: 38

Sequence: 1 AAADPHE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	600	2 A47391	serum albumin prec
2	38	100.0	609	1 ABHUS	serum albumin prec
3	34	89.5	177	2 D71311	hypothetical prote
4	34	89.5	575	2 C70781	probable fprB prot
5	34	89.5	740	2 E69420	hydrogenase expres
6	33	86.8	238	2 H75392	hypothetical prote
7	33	86.8	308	2 F86147	hypothetical prote
8	33	86.8	345	2 AH3101	IS5 family transpo
9	33	86.8	351	2 B98185	hypothetical prote
10	33	86.8	414	1 C70859	probable hexosyltr
11	33	86.8	438	2 E87123	probable transfera
12	32	84.2	363	2 A95504	X-pro aminopeptida
13	32	84.2	421	2 AH3200	hydroxypyruvate re
14	31	81.6	105	2 C72631	hypothetical prote
15	31	81.6	238	1 S72623	ribulose-5-phospha
16	31	81.6	263	2 D71281	probable pyroline
17	31	81.6	272	2 T32959	hypothetical prote
18	31	81.6	300	2 T34213	hypothetical prote
19	31	81.6	345	2 T25893	hypothetical prote
20	31	81.6	418	2 A96601	hypothetical prote
21	31	81.6	634	2 T49415	hypothetical prote
22	31	81.6	713	2 A28706	calpastatin, cardi
23	31	81.6	883	2 A84210	hypothetical prote
24	31	81.6	1316	2 D87145	[beta], subunit of
25	31	81.6	1316	2 S31146	DNA-directed RNA p
26	31	81.6	1316	2 G70535	probable rpoC prot
27	31	81.6	1863	1 A58881	breat/ovarian can
28	30	78.9	152	2 C75544	hypothetical prote
29	30	78.9	232	2 G70738	probable PPE prote

ALIGNMENTS

RESULT 1

A47391

serum albumin precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)

C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999

C:Accession: A47391

R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwolet, J.; Putnam, F

Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993

A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilir

A:Reference number: A47391; MUID:93211971; PMID:8460152

A:Contents: B/B homozygote

A:Accession: A47391

A>Status: preliminary

A:Molecule type: mRNA, protein

A:Residues: 1-600 <WAT>

A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295

A:Experimental source: liver

A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:P:128281)

C:Superfamily: serum albumin; serum albumin repeat homology

F:21-194/Domain: serum albumin repeat homology <SA1>

F:213-386/Domain: serum albumin repeat homology <SA2>

F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 38; DB 2; Length 600;

Best Local Similarity 100.0%; Pred. No. 4.1; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0;

QY 1 AAADPHE 7

DB 378 AAADPHE 384

RESULT 2

ABHUS

serum albumin precursor [validated] - human

N:Alternate names: preproalbumin

N:Contains: kinetensin

C:Species: Homo sapiens (man)

C:Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text_change 17-Mar-2000

C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S31

R:Law, R.M.; Aaelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu:

Nucleic Acids Res. 9, 6103-6114, 1981

A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia col:

A:Reference number: A93743; MUID:82081882; PMID:6171778

A:Accession: A93743

A:Molecule type: mRNA

A:Residues: 1-419, 'K' 421-609 <LAW>

A:Cross-references: EMBL:V00495; GB:J000078; GB:L00132; NID:G285591; PIDN:CAA2:

R:Dugaiczyk, A.; Law, S.W.; dennison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A:Reference number: A93936; MUID:82105994; PMID:6275391

hypothetical prote
hypothetical prote
probable transcript
hypothetical prote
conserved hypotet
phosphonate monos
adenine deaminase
adenine deaminase
ferredoxin-nitrite
adenine deaminase
pyruvate carboxyla
valine-tRNA ligase
hypothetical prote
conserved hypotet
DNA polymerase III
CPY protein - midg

A;Accession: A93936
A;Molecule type: mRNA
A;Residues: 1-120, 'G', 122-609 <DUG>
A;Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590
R;Uranio, Y.; Watanabe, K.; Sakai, M.; Tamaki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A;Reference number: I39427; MUID:86140099; PMID:2419329
A;Accession: I39427
A;Status: translation not shown
A;Molecule type: protein
A;Residues: 1-26 <URA>
A;Cross-references: GB:M13075; NID:G178330; PIDN:AA51688.1; PID:G553173
R;Watkins, S.; Madison, J.; Galliano, M.; Minichioti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.
A;Reference number: I59286; MUID:94181575; PMID:8134387
A;Accession: I59286
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 282-290, 'KSRFDLQ' <WAT>
A;Cross-references: GB:S69192; NID:G546032; PIDN:AA30282.1; PID:G546033
A;Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia
R;Madison, J.; Galliano, M.; Watkins, S.; Minichioti, L.; Porta, F.; Rossi, A.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-
A;Reference number: I59313; MUID:94294404; PMID:8022807
A;Accession: I59313
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 589-590, 'ALPRVKNLLQVLP' <MAD>
A;Cross-references: GB:S70799; NID:G547231; PIDN:AA31177.1; PID:G547232
A;Note: This frame-shift variant is designated albumin Bazzano; four additional variants
R;Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A;Reference number: G01747
A;Accession: G01747
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-120, 'G', 122-455 <VEN>
A;Cross-references: EMBL:U22361; NID:G763428; PIDN:AAA64922.1; PID:G763431
R;Meloun, B.; Moravsek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A;Title: Complete amino acid sequence of human serum albumin.
A;Reference number: A91420; MUID:76187907; PMID:1225573
A;Accession: A91420
A;Molecule type: protein
A;Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-
R;Roehr, U.; Spittler, G.; Tripiet, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from u
A;Reference number: S06422
A;Note: this paper is in German, with an English abstract
A;Accession: S06422
A;Molecule type: protein
A;Residues: 25-48 <ROS>
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A;Title: Mass spectrometric identification of modifications to human serum albumin treat
A;Reference number: S36882; MUID:93384321; PMID:8373198
A;Accession: S36882
A;Molecule type: protein
A;Residues: 45-67; 141-160; 311-337; 469-490; 570-581 <FIN>
R;Kausler, E.; Spittler, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A;Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol

A;Reference number: S17599; MUID:92126241; PMID:1772598
A;Accession: S17599
A;Molecule type: protein
A;Residues: 25-54; 354-357; 431-447 <KAU>
A;Note: 49-leu was also found
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid protea
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: A45800
A;Molecule type: protein
A;Residues: 166-173 <CAR>
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; W
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tr
A;Reference number: A03239; MUID:86242180; PMID:3087352
A;Accession: A03239
A;Molecule type: protein
A;Residues: 166-173, 'L' <MOG>
R;Galliano, M.; Minichioti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, J
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A;Title: Mutations in genetic variants of human serum albumin found in Italy.
A;Reference number: A38255; MUID:91062352; PMID:2247440
A;Accession: C38255
A;Molecule type: protein
A;Residues: 76-111 <GALL>
A;Accession: B38255
A;Molecule type: protein
A;Residues: 82-105, 'K', 107-110 <GALL>
A;Note: this variant is designated albumin Vibo Valentia
A;Accession: A38255
A;Molecule type: protein
A;Residues: 76-83, 'K', 85-106 <GALL>
A;Note: this variant is designated albumin Torino
R;Minichioti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin He
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: protein
A;Residues: 255-263, 'E', 265-281 <MIN1>
A;Note: this variant is designated albumin Herborn
R;Minichioti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta, F
Biochim. Biophys. Acta 1119, 232-238, 1992
A;Title: Two albumins with identical electrophoretic mobility are produced by diffe
A;Reference number: S21078; MUID:92190239; PMID:1347703
A;Accession: S21078
A;Molecule type: protein
A;Residues: 354-356, 'K', 358-378 <MIN2>
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported
R;He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A;Title: Atomic structure and chemistry of human serum albumin.
A;Reference number: A46756; MUID:92334427; PMID:1630489
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.
in: The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,
A;Reference number: A94442
A;Contents: annotation; three-dimensional structure and disulfide bonds
R;Sabat, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A;Title: Disulfide bonds in human serum albumin.
A;Reference number: A90930
A;Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A;Reference number: A90299; MUID:78186630; PMID:656055
A;Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
in: Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20,
A;Title: Serum albumin: conformation and active sites.
A;Reference number: A94408

QY 1 AAADPHE 7
|||
Db 85 AAVDPHE 91
|||
RESULT 4
C70781
probable fprB protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Mar-2000
C/Accession: C70781
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: C70781
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-575 <COL>
A/Cross-references: GB:Z731101; GB:AL123456; NID:G3261565; PIDN:CAA97393.1; PID:e241991;
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: fprB
C/Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h
F;2-64/Domain: ferredoxin 2[4Fe-4S] homology <PER>
Query Match 89.5%; Score 34; DB 2; Length 575;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 AADPHE 7
|||
Db 270 AADPHE 275
|||
RESULT 5
E69420
hydrogenase expression/formation regulatory protein (hypF) homolog - Archaeoglobus fulg
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 14-Apr-2003
C/Accession: E69420
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Uutterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: E69420
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-740 <MLE>
A/Cross-references: GB:AE001009; GB:AE000782; NID:G2689332; PIDN:AA89876.1; PID:G264921
C/Superfamily: carbamoyl phosphate-converting enzyme ([NiFe])-hydrogenase maturation fac
Query Match 89.5%; Score 34; DB 2; Length 740;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAADPHE 7
|||
Db 606 AAADPHE 612
|||
RESULT 6
H75392
hypothetical protein - Deinococcus radiodurans (strain R1)
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

RESULT 6
H75392
hypothetical protein - Deinococcus radiodurans (strain R1)
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 31-Mar-2000

C:Accession: H75392
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567286
 A:Accession: H75392
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-238 <WHI>
 A:Cross-references: GB:AE001590; GB:AE000513; NID:g6459214; PIDN:AAF11027.1; PID:g6459214
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1454
 A:Map position: 1

Query Match 86.8%; Score 33; DB 2; Length 238;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPH 6
 |||||
 DB 162 AAADPH 167

RESULT 7

hypothetical protein TUN6.5 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C:Accession: F86147
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: F86147
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-308 <STO>
 A:Cross-references: GB:AE005172; NID:g8671835; PIDN:AAF78398.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 96.8%; Score 33; DB 2; Length 308;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPH 6
 |||||
 DB 117 AAADPH 122

RESULT 8

IS5 family transposase tnp [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AH3101
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
 Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AH3101
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-345 <KUR>
 A:Cross-references: GB:AE008689; PIDN:AA145230.1; PID:gl7742912; GSPDB:GN00187
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: tnp
 A:Map position: linear chromosome

Query Match 96.8%; Score 33; DB 2; Length 345;
 Best Local Similarity 95.7%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAADPH 7
 |||||
 DB 167 ALADPH 173

RESULT 9

B98185
 hypothetical protein AGR_L_863gl [imported] - Agrobacterium tumefaciens (strain C58, Ce
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 C:Accession: B98185
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourollo, B.; Goldman
 A.; Liu, P.; Wollam, C.; Allinger, D.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
 A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: B98185
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-351 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK89004.1; PID:gl5158792; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L_863gl
 A:Map position: linear chromosome

Query Match 96.8%; Score 33; DB 2; Length 351;
 Best Local Similarity 95.7%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAADPH 7
 |||||
 DB 173 ALADPH 179

RESULT 10

C70859
 Probable hexosyltransferase (EC 2.4.1.-) Rv3032 [similarity] - Mycobacterium tuberculosis
 C:Species: Mycobacterium tuberculosis
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 16-Jun-2000
 C:Accession: C70859
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: C70859
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-414 <COL>
 A:Cross-references: GB:AL021287; GB:AL123456; NID:g3261508; PIDN:CAA16117.1; PID:g27916
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv3032

C:Superfamily: probable hexosyltransferase ytxN
 C:Keywords: Glycosyltransferase; hexosyltransferase

A;Residues: 1-238 <JUH>
A;Cross-references: EMBL:X83571; NID:G609673; PIDN:CAA8554.1; PID:G609674
R;Juhnke, H.
submitted to the EMBL Data Library, December 1994
A;Reference number: S51587
A;Accession: S51587
A;Molecule type: DNA
A;Residues: 1-238 <JUZ>
A;Cross-references: EMBL:X83571; NID:G609673; PIDN:CAA8554.1; PID:G609674
R;Cziepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S56891
A;Accession: S56902
A;Molecule type: DNA
A;Residues: 1-238 <CZI>
A;Cross-references: EMBL:Z49396; NID:G1008312; PIDN:CAA89415.1; PID:G1008313; GSPDB:GN00
C;Genetics:
A;Gene: SGD:RPE1; POS18; SGD:S0003657
A;Cross-references: MIPS:YJL121c; SGD:S0003657
A;Map position: 10L
C;Superfamily: yeast ribulose-5-phosphate-epimerase
C;Keywords: carbohydrate metabolism; isomerase; leucine zipper

Query Match 81.6%; Score 31; DB 1; Length 238;
Best Local Similarity 83.3%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AADPHE 7
DB 213 AADPHD 218
|||||:

Search completed: April 19, 2004, 13:02:30
Job time : 1.85319 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.510619 Seconds

(without alignments)

713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_362_368

Perfect score: 38

Sequence: 1 AADPHE 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	600	1 ALBU_MACMU	Q28522 macaca mula
2	38	100.0	609	1 ALBU_HUMAN	P02768 homo sapien
3	34	89.5	177	1 Y552_TREPA	O83563 treponema p
4	34	89.5	575	1 FRAB_MYCTU	Q10547 mycobacteri
5	31	81.6	238	1 RPE_YEAST	P46969 saccharomyc
6	31	81.6	263	1 PROC_TREPA	P27771 treponema p
7	31	81.6	350	1 G3P_METKA	P58839 methanopyru
8	31	81.6	504	1 PRDB_HUMAN	Q95q75 homo sapien
9	31	81.6	713	1 ICAL_PIG	P12675 sus scrofa
10	31	81.6	742	1 PKWA_TRECU	P49695 thermomonos
11	31	81.6	1316	1 RPOC_MYCLE	P30761 mycobacteri
12	31	81.6	1316	1 RPOC_MYCTU	P47769 mycobacteri
13	31	81.6	1863	1 BRC1_HUMAN	P38398 homo sapien
14	31	81.6	1863	1 BRC1_PANTR	Q9GKx8 pan troglod
15	30	78.9	232	1 Y726_MYCTU	Q50702 mycobacteri
16	30	78.9	393	1 Y752_TREPA	O83743 treponema p
17	30	78.9	455	1 TBB_DICDI	P32256 dictyosteli
18	30	78.9	538	1 ADESC_METH	O28952 methanobact
19	30	78.9	556	1 ADESC_ARCFU	O29999 archaeoglob
20	30	78.9	569	1 NIR_MAIZE	P17847 zea mays (m
21	30	78.9	877	1 SYV_METH	O26861 methanobact
22	30	78.9	887	1 FTSK_RHILO	Q98eh3 rhizobium l
23	30	78.9	922	1 GVEA_AERSA	P48369 aeromonas s
24	30	78.9	1433	1 DP03_BACHD	Q9ka72 bacillus ha
25	29	76.3	130	1 ODOI_HOUSE	Q60597 mus musculu
26	29	76.3	144	1 G10_XENLA	P12805 xenopus lae
27	29	76.3	146	1 HBB_CRONI	P02129 crocodylus
28	29	76.3	194	1 PUR1_LACCA	P35853 lactobacill
29	29	76.3	209	1 AT13_VACCV	P21117 vaccinia vi
30	29	76.3	314	1 SIX3_CHICK	O42406 gallus gall
31	29	76.3	317	1 UL14_HSVB	P28949 equine herp
32	29	76.3	322	1 AT12_VACCC	P21114 vaccinia vi
33	29	76.3	344	1 HIFN_BRARE	P59723 brachydanio

ALIGNMENTS

RESULT 1

ID	ALBU_MACMU	STANDARD	PRT	600 AA
AC	Q28522			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Serum albumin precursor (fragment).			
GN	ALB			
OS	Macaca mulatta (Rhesus macaque).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;			
OC	Cercopitheidae; Macaca.			
OX	NCBI_TaxID=9544;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93211971; PubMed=8460152;			
RA	Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,			
RA	Dwulet J., Putnam F.W.;			
RT	"cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding."			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).			
CC	-1- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- TISSUE SPECIFICITY: Plasma.			
CC	-1- SIMILARITY: Belongs to the ALB/APP/VDB family.			
CC	-1- SIMILARITY: Contains 3 albumin domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).			
DR	EMBL: M90463; AAA36906.1; -			
DR	PIR: A47391; A47391.			
DR	HSP: P02768; I57B.			
DR	InterPro: IPR000264; Serum_albumin.			
DR	Pfam: PF00273; transport_prot; 3.			
DR	PRINTS: PR00802; SERUMALBUMIN.			
DR	ProDom: PD002486; Serum_albumin; 1.			
DR	SMART: SM00103; ALBUMIN; 3.			
DR	PROSITE: PS00212; ALBUMIN; 3.			
KW	Metal-binding; lipid-binding; Repeat; Signal; Copper.			
FT	NON TER	1		
FT	SIGNAL	<1	10	BY SIMILARITY.
FT	PROPEP	11	16	BY SIMILARITY.
FT	CHAIN	17	600	SERUM ALBUMIN.
FT	DOMAIN	17	197	ALBUMIN 1.
FT	DOMAIN	204	389	ALBUMIN 2.
FT	DOMAIN	396	587	ALBUMIN 3.

Q50727 m probable
P41385 bombyx mori
Q8p775 xanthomonas
Q95078 drosophila
O35681 mus musculu
P40748 rattus norv
Q9bqg1 homo sapien
P05314 spinacia ol
Q9ph47 xyiella fas
Q8rfe63 xyiella fas
P40991 saccharomyc
P38769 saccharomyc

FT METAL 19 COPPER (BY SIMILARITY).
FT BINDING 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 BY SIMILARITY.
FT DISULFID 91 BY SIMILARITY.
FT DISULFID 106 BY SIMILARITY.
FT DISULFID 140 BY SIMILARITY.
FT DISULFID 184 BY SIMILARITY.
FT DISULFID 216 BY SIMILARITY.
FT DISULFID 261 BY SIMILARITY.
FT DISULFID 281 BY SIMILARITY.
FT DISULFID 294 BY SIMILARITY.
FT DISULFID 332 BY SIMILARITY.
FT DISULFID 377 BY SIMILARITY.
FT DISULFID 385 BY SIMILARITY.
FT DISULFID 408 BY SIMILARITY.
FT DISULFID 453 BY SIMILARITY.
FT DISULFID 477 BY SIMILARITY.
FT DISULFID 493 BY SIMILARITY.
FT DISULFID 503 BY SIMILARITY.
FT DISULFID 530 BY SIMILARITY.
FT DISULFID 575 BY SIMILARITY.
FT DISULFID 574 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;
Query Match 100.0%; Score 38; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 2.6; 0; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;
QY 1 AAADPHE 7
Db 378 AAADPHE 384
RESULT 2
ALBU HUMAN
ID ALBU HUMAN STANDARD; PRT; 609 AA.
AC P02768; O95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] SEQUENCE FROM N.A.
RP SEQUENCE 86196112; PubMed=3009475;
RX MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczky A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX MEDLINE=82081882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wilson K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczky A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).

RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver.";
RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
RA Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin.";
RN Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marudan K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carrinchi P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers K.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RN FEBS Lett. 58:134-137(1975).
RN [9]
RP SEQUENCE OF 25-609.
RA Brown J.R., Shockley P., Behrens P.O.;
RL (in) Bing D.H. (eds.);
RL The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RP SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Menaya J., Parrilla R., Ayuso M.S.;
RN Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RP SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RP SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994.";

RL Electrophoresis 15:1459-1465(1994).
RN [14]
RP DISULFIDE BONDS.
RX Saber M.A., Stockbauer P., Moravsek L., Meloun B.;
RA "Disulfide bonds in human serum albumin."
RT Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15]
RP BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobson C.;
RT "Lysine residue 240 of human serum albumin is involved in high-affinity binding of bilirubin."
RN Biochem. J. 171:453-459(1978).
RN [16]
RP VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second domain of serum albumin."
RN Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RP VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=86068523; PubMed=3479777;
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M., Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from Amerindian and Japanese populations."
RN Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18]
RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huss K., Ishio K., Satoh C., Fujita M., Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese allolalbumins."
RN Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19]
RP VARIANTS MANAUS; OSAKA; NAGOVA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia."
RN Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20]
RP DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of human serum albumin whose precursor has an aberrant signal peptidase cleavage site."
RN Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21]
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferris G., Madison J., Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in Italy."
RN Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22]
RP VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M., Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two carboxyl-terminal variants of human serum albumin."
RN Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23]
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOWAGOME-3 HIS-23; KOWAGOME-2 ARG-152 AND KOWAGOME-1 GLU-396.
RX MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E., Matsuda Y.-I., Amaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese."
RN Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).

RN [24]
RP VARIANT CASEBROOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum albumin: albumin Casebrook (494 Asp-->Asn)."
RN Biochim. Biophys. Acta 1097:49-54(1991).
RN [25]
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190339; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferris G., Crespeau H., Rochu D., Porta F.;
RT "Two allolalbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 38; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 2.7; Mismatches 0; Gaps 0;
Matches 7; Conservative 0; Indels 0; Gaps 0;
QY 1 AAADPHE 7
DB 386 AAADPHE 392

RESULT 3
Y552_TREPA
ID Y552_TREPA STANDARD; PRT; 177 AA.
AC O83563;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0552.
GN TP0552.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G., Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Usterback T., McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S., Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis spirochete."
RN Science 281:375-388(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
DR EMBL; AE001230; AAC65538.1; -.
DR PIR; D71311; D71311.
DR TIGR; TP0552; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 177 AA; 20172 MW; 59A0F0B815D006EE CRC64;
Query Match 89.5%; Score 34; DB 1; Length 177;
Best Local Similarity 85.7%; Pred. No. 4.9;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAADPHE 7
DB 85 AAADPHE 91

RESULT 4
ID_FPRB MYCTU STANDARD; PRT; 575 AA.
AC Q10547;
DT 01-OCT-1996 (Rel. 34, Created)
DI 01-OCT-1996 (Rel. 34, Last sequence update)
DE Probable ferredoxin/ferredoxin-NADP reductase (EC 1.18.1.2) (FNR).
GN FPRB OR RV0886 OR MT0909 OR MT0331.14 OR MB0910.
OS Mycobacterium tuberculosis, and
OC Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala P.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh J., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RA MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Wadman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES=M.bovis; STRAIN=AF2122/97;
RA MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duboy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- CATALYTIC ACTIVITY: Reduced ferredoxin + NADP(+) = oxidized
CC ferredoxin + NADPH.
CC -!- COFACTOR: FAD; probably binds one or two 4Fe-4S clusters.
CC -!- SIMILARITY: In the N-terminal, belongs to the bacterial-type
CC ferredoxin family.
CC -!- SIMILARITY: IN THE C-TERMINAL, TO OTHER FERREDOXIN NADP
CC REDUCTASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 273101; CAA97393.1; -
CC DR EMBL; AE006978; AAK45151.1; -
CC DR EMBL; BX248336; CAD93771.1; -
CC DR FIB; C70781; C70781.
CC DR HSP; Q45560; 1BD6.
CC TIGR; MT0909; -

Tuberculin; RV0886; -
DR InterPro; IPR001450; 4Fe4S ferredoxin.
DR InterPro; IPR000813; 7Fe ferredoxin.
DR InterPro; IPR000759; Adnrdx reductase.
DR InterPro; IPR001327; FAD_PyT_redox.
DR Pfam; PF00037; fer4; 2.
DR PRINTS; PR00354; 7FESFRDOXIN.
DR PRINTS; PR00419; ADNRDASE.
DR PRINTS; PR00368; FADPNR.
DR PROSITE; PS00198; 4Fe4S FERREDOXIN; 1.
DR PROSITE; PS00198; 4Fe4S FERREDOXIN; 1.
KW Oxidoreductase; Flavoprotein; NADP; FAD; Electron transport;
KW Iron-sulfur; 4Fe-4S; Complete proteome.
FT DOMAIN 1 83 FERREDOXIN.
FT DOMAIN 115 575 FERREDOXIN-NADP REDUCTASE.
FT METAL 15 9 IRON-SULFUR 1 (BY SIMILARITY).
FT METAL 15 15 IRON-SULFUR 1 (BY SIMILARITY).
FT METAL 19 19 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 46 46 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 49 49 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 52 52 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 56 56 IRON-SULFUR 1 (BY SIMILARITY).
SQ SEQUENCE 575 AA; 61337 MW; 170C58599872A810 CRC64;
Query Match 89.5%; Score 34; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 AADPHE 7
Db 270 AADPHE 275
RESULT 5
RPE_YEAST STANDARD; PRT; 238 AA.
ID_RPE YEAST
AC P46569;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ribulose-phosphate 3-epimerase (EC 5.1.3.1) (Pentose-5-phosphate 3-
DE epimerase) (PPE) (RPE).
GN RPE1 OR POS18 OR XLI21C OR J0731.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetaceae;
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Junke H.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=97103775; PubMed=8948101;
RA Cziepluch C., Kordes B., Pujol A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT reveals 19 open reading frames including URA2 [5' end], TRK1, PBS2,
RT SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CCT7, GZF3, two tRNA genes,
RT three remnant delta elements and a Ty4 transposon."
RL Yeast 12:1471-1474(1996).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=97074231; PubMed=8929392;
RA Miosga T., Zimmermann F.K.;
RT "Cloning and characterization of the first two genes of the non-
RT oxidative part of the Saccharomyces cerevisiae pentose-phosphate
RT pathway."
RL Curr. Genet. 30:404-409(1996).
CC -!- CATALYTIC ACTIVITY: D-ribulose 5-phosphate = D-xylose 5-
CC phosphate.
CC -!- PATHWAY: Nonoxidative branch of the pentose phosphate pathway.
CC -!- SIMILARITY: Belongs to the ribulose-phosphate 3-epimerase family.
CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

EMBL; X83571; CA858534.1; -.
EMBL; Z49396; CA89415.1; -.
PIR; S72623; S72623.
HSSP; Q43843; 1RPX.
GermOnline; 141733; -.
SGD; S0003657; RPE1.
GO; GO:0004750; F:Ribulose-phosphate 3-epimerase activity; IDA.
GO; GO:0006098; P:pentose-phosphate shunt; IMP.
InterPro; IPR000056; Ribul_P_3_epim.
Pfam; PF00834; Ribul_P_3_epim; 1.
TIGRfam; TIGR01163; rpe; 1.
PROSITE; PS01085; RIBUL_P_3_EPIMER_1; 1.
PROSITE; PS01086; RIBUL_P_3_EPIMER_2; 1.
Isonerases; Carbohydrate metabolism.
DOMAIN 64 75 COULD BE PART OF AN INTRONIC SEQUENCE.
SEQUENCE 238 AA; 25967 MW; 7EEFASBEID3DF8F CRC64;
SF

```

Query Match	81.6%;	Score 31;	DB 1;	Length 238;
Best Local Similarity	83.3%;	Pred. No. 28;		
Matches 5;	Conservative 1;	Mismatches	0;	Indels 0;
		Gaps	0;	

QY 2 AADPHE 7
 ||||:
 Db 213 AADPHD 218

RESULT 6	PROC TREPA	STANDARD;	PRT;	263 AA.
ID	PROC TREPA			
AC	P27771; O83775;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Pyroline-5-carboxylate reductase (BC 1.5.1.2) (P5CR) (P5C reductase).			
GN	PROC OR TP0797.			
OS	Treponema pallidum.			
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.			
OX	NCBI_TaxID=160;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Nichols;			
RA	MEDLINE=90264287; PubMed=218947;			
RX	Gherardini F.C., Hobbs M.M., Stamm L.V., Bassford P.J. Jr.;			
RT	"Complementation of an Escherichia coli proC mutation by a gene cloned			
RT	from Treponema pallidum."			
RL	J. Bacteriol. 172:2996-3002(1990).			
RN	[2]			

RP SEQUENCE FROM N.A.
 RT STRAIN=Nichols;
 RC MEDLINE=9832770; PubMed=9665876;
 RX Fraser C.W., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg M., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
 RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of *Treponema pallidum*, the syphilis
 spirochete.";
 RL Science 281:375-388(1998).

CC	- - CATALYTIC ACTIVITY: L-proline + NAD(P) ⁺ = 1-pyrroline-5-
CC	carboxylate + NAD(P)H.
CC	- - PATHWAY: proline biosynthesis; third (last) step.
CC	- - SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC	- - SIMILARITY: Belongs to the pyrroline-5-carboxylate reductase

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AE010355; AAM01833.1; --
DR HAMAB; MF 00559; -- 1.
DR InterPro; IPR000173; GAP_dhhydrogenase.
DR InterPro; IPR000436; GAPDH-II_archae.
DR Pfam; PF00044; gpdh; 2.
DR Pfam; PF02800; gpdh; 1.
DR ProDom; PD007761; GAPDH-II_archae; 1.
DR TIGRFAMs; TIGR01546; GAPDH-II_archae; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NADP; Complete proteome.
FT BINDING 148 148 GLYCERALDEHYDE 3-PHOSPHATE (BY
SIMILARITY).
SQ SEQUENCE 350 AA; 38367 MW; F9BCB72702E22E02 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 350;
Best Local Similarity 83.3%; Pred. No. 42;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AADPHE 7
|:|:|:|:
DB 177 AADPHQ 182

RESULT 8

PRDB HUMAN
ID PRDB HUMAN STANDARD; PRT; 504 AA.
AC Q9NQV5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE PR-domain protein 11.
GN PRDM11 OR PF08.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang X.-H., Huang S.;
RT "A family of novel PR-domain (PRDM) genes as candidate tumor suppressors";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 SET domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF275818; AAF87244.1; --
DR Genew; HGNC:12996; PRDM11.
DR InterPro; IPR001214; SET.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS0280; SET; 1.
FT DOMAIN 142 257 SET.
SQ SEQUENCE 504 AA; 57032 MW; CF591094CCF45515 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 504;
Best Local Similarity 83.3%; Pred. No. 61;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AADPHE 7
|:|:|:|:
DB 426 ASDPHE 431

RESULT 9

ICAL PIG
ID ICAL PIG STANDARD; PRT; 713 AA.
AC P12675;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calpain inhibitor (Calpastatin).
GN CAST.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=88241021; PubMed=2837276;
RA Takano E., Maki M., Mori H., Hatanaka M., Marti T., Titani K.,
RA Kannagi R., Marti T., Titani K., Murachi T.;
RT "Pig heart calpastatin: identification of repetitive domain structures and anomalous behavior in polyacrylamide gel electrophoresis";
RL Biochemistry 27:1964-1972(1988).
RN [2]
RP SEQUENCE OF 304-554 FROM N.A.
RX MEDLINE=87054580; PubMed=3780962;
RA Takano E., Maki M., Hatanaka M., Mori H., Zenita K., Sakiyama T.,
RA Kannagi R., Marti T., Titani K., Murachi T.;
RT "Evidence for the repetitive domain structure of pig calpastatin as demonstrated by cloning of complementary DNA";
RL FEBS Lett. 208:199-202(1986).
CC -!- FUNCTION: Specific inhibition of calpain (calcium-dependent cysteine protease). Plays a key role in postmortem tenderization of meat and have been proposed to be involved in muscle protein degradation in living tissue.
CC -!- DOMAIN: Has four inhibitory domains.
CC -!- SIMILARITY: belongs to the calpastatin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M20160; AAA31012.1; --
DR EMBL; M27969; AAA31003.1; --
DR PIR; A24627; A24627.
DR PIR; A28706; A28706.
DR InterPro; IPR001259; Calpain_inhib.
DR Pfam; PF00748; Calpain_inhib; 4.
KW Repeat; Thiol protease inhibitor.
FT REPEAT 171 223 INHIBITORY DOMAIN 1.
FT REPEAT 307 359 INHIBITORY DOMAIN 2.
FT REPEAT 447 500 INHIBITORY DOMAIN 3.
FT REPEAT 583 636 INHIBITORY DOMAIN 4.
FT CONFLICT 328 328 L > V (IN REF. 2).
SQ SEQUENCE 713 AA; 77123 MW; ASD4E8F119CE97B5 CRC64;

Query Match 81.6%; Score 31; DB 1; Length 713;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AADPHE 7
|:|:|:|:
DB 273 AADPHE 279

RESULT 10

PKWA THECU
ID PKWA_THECU STANDARD; PRT; 742 AA.

```
P49695;
01-FEB-1996 (Rel. 33, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable serine/threonine-protein kinase pKWA (EC 2.7.1.37).
GN PKWA OR PKW1.
OS Thermomonospora curvata.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptosporangineae; Thermomonosporaceae; Thermomonospora.
OX NCBI_TaxID=2020;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 3352;
RX MEDLINE=96200125; PubMed=8631732;
RA Janda L., Tichy P., Spizek J., Petrcek M.;
RT "A deduced Thermomonospora curvata protein containing
RL serine/threonine protein kinase and WD-repeat domains.";
CC J. Bacteriol. 178:1487-1489(1996).
CC -!- FUNCTION: MAY PLAY A REGULATORY ROLE DURING THE COMPLEX GROWTH
CC CYCLE AND IN SECONDARY METABOLITE PRODUCTION.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC -!- SIMILARITY: Contains 7 WD repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF115313; AAB05922.1;
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR ProDom; PD000016; WD40; 5.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 5.
DR PROSITE; PS00082; WD_REPEATS_2; 7.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW WD repeat.
FT DOMAIN 16 266 PROTEIN KINASE.
FT NP BIND 22 30 ATP (BY SIMILARITY).
FT BINDING 44 44 ATP (BY SIMILARITY).
FT ACT_SITE 138 138 BY SIMILARITY.
FT DOMAIN 301 394 PRO/GLU/ASP-RICH (SPACER).
FT REPEAT 455 496 WD 1.
FT REPEAT 497 538 WD 2.
FT REPEAT 539 580 WD 3.
FT REPEAT 581 621 WD 4.
FT REPEAT 622 653 WD 5.
FT REPEAT 654 705 WD 6.
FT REPEAT 706 742 WD 7.
SQ SEQUENCE 742 AA; 78950 MW; ACL734640DB4383D CRC64;

Query Match
Best Local Similarity 81.6%; Score 31; DB 1; Length 742;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AADPHE 7
:|||||
DB 444 SADPHE 449
```

```
RESULT 11
RPOC MYCLE
ID RPOC MYCLE STANDARD; PRT; 1316 AA.
AC P30761.
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
DE beta' chain) (RNA polymerase beta' subunit).
GN RPOC OR MLI1890.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93188701; PubMed=8446028;
RA Honore N.T., Bergh S., Chanteau S., Doucet-Populaire F.,
RA Eigmeier K., Garnier T., Georges C., Launois P., Limpaboon T.,
RA Newton S., Niang K., del Portillo P., Ramesh G.R., Reddi P.,
RA Ridet P.R., Sittisombut N., Wu-Hunter S., Cole S.T.;
RT "Nucleotide sequence of the first cosmid from the Mycobacterium
RT leprae genome project: structure and function of the Rif-Str
RT regions.";
RL Mol. Microbiol. 7:207-214(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltingworth T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- SUBUNIT: The enzyme consists of the sigma chain and the core
CC enzyme which is composed of 2 alpha chains, 1 beta chain, and 1
CC beta' chain.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z14314; CAA78669.1; -.
DR EMBL; AL583923; CAC30844.1; -.
DR PIR; D87145; D87145.
DR PIR; S31146; S31146.
DR HSP; Q9KWU6; IHQM.
DR Leproma; MLI1890; -.
DR InterPro; IPR000722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_polA_N.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF00623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
```

DR Pfam: PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam: PF04998; RNA_pol_Rpb1_5; 1.
DR SMART: SM00663; RPOA_N; 1.
RW Transferase; DNA-directed RNA polymerase; Transcription;
KW Complete proteome.
FT CONFLICT 598 A -> R (IN REF. 1).
SQ SEQUENCE 1316 AA; 146895 MW; 7131AB63B53F6BCF CRC64;

Query Match 81.6%; Score 31; DB 1; Length 1316;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AADPHE 7
Db 1115 SADPHE 1120
:|||||

RESULT 12
RPOC MYCTU STANDARD; PRT; 1316 AA.
AC P47769; O06771;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
DE beta' chain) (RNA polymerase beta' subunit).
GN RPOC OR RV0668 OR MT0696 OR MTC1376.07C OR MB0687.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Barry C.E. III, Tekala F.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Connor R.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jajelski K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Biethal W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490(2002).
RN [3]
RP SEQUENCE OF 1-148 FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37RV;
RX MEDLINE=94304130; PubMed=8031050;
RA Miller L.P., Crawford J.T., Shinnick T.M.;
RT "The rpoB Gene of Mycobacterium tuberculosis."
RL Antimicrob. Agents Chemother. 38:805-811(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF122/97;
RX MEDLINE=22709107; PubMed=12789972;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Dutthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: The enzyme consists of the sigma chain and the core
CC enzyme which is composed of 2 alpha chains, 1 beta chain, and 1
CC beta' chain.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta' chain family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: Z95972; CAB09389.1; -.
CC EMBL: AE006964; AAK44922.1; -.
CC EMBL: L27989; AAK21417.1; -.
CC EMBL: BX248336; CAD93549.1; -.
CC PIR: G70535; G70535.
CC HSSP: Q9KW06; IHQM.
CC TIGR: MT0696; -.
CC TubercuList; RV0668; -.
CC InterPro: IPR007072; RNA_pol_A.
CC InterPro: IPR007080; RNA_pol_Rpb1_1.
CC InterPro: IPR007066; RNA_pol_Rpb1_3.
CC InterPro: IPR007083; RNA_pol_Rpb1_4.
CC InterPro: IPR007081; RNA_pol_Rpb1_5.
CC InterPro: IPR006592; RNA_pol_A_N.
CC Pfam: PF04997; RNA_pol_Rpb1_1; 1.
CC Pfam: PF00623; RNA_pol_Rpb1_2; 1.
CC Pfam: PF04983; RNA_pol_Rpb1_3; 1.
CC Pfam: PF05000; RNA_pol_Rpb1_4; 1.
CC Pfam: PF04998; RNA_pol_Rpb1_5; 1.
CC SMART: SM00663; RPOA_N.1
KW Transferase; DNA-directed RNA polymerase; Transcription;
KW Complete proteome.
FT CONFLICT 594 594 G -> E (IN REF. 2).
SQ SEQUENCE 1316 AA; 146769 MW; 45BF24839AF53E8B CRC64;

Query Match 81.6%; Score 31; DB 1; Length 1316;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AADPHE 7
Db 1115 SADPHE 1120
:|||||

RESULT 13
BRCL_HUMAN STANDARD; PRT; 1863 AA.
ID BRCL_HUMAN
AC P38398;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Breast cancer type 1 susceptibility protein.
GN BRCA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT ARG-1775.
RX MEDLINE=95025996; PubMed=7545954;
RA Miki Y., Swensen J., Shattuck-Bidens D., Futreal P.A., Harshman K.,
RA Tavtigian S., Liu Q., Cochran C., Bennett L.M., Ding W., Bell R.,

RA Rosenthal J., Husey C., Tran T., McClure M., Frye C., Hattier T.,
RA Shells R., Haugen-Strano A., Kacher H., Yakumo K., Gholami Z.,
RA Shaffer D., Stone S., Bayer S., Wray C., Bogden R., Dayananth P.,
RA Ward J., Tonin P., Narod S., Bristow P.K., Norris F.H., Helvaring L.,
RA Morrison P., Rostek P., Lai M., Barrett J.C., Lewis C., Neuhausen S.,
RA Cannon-Albright L., Godlgr D., Wiseman R., Kamb A., Skolnick M.H.,
RT "A strong candidate for the breast and ovarian cancer susceptibility
RT gene BRCA1".
RL Science 266:66-71(1994).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=97092865; PubMed=8939427;
RX Smith T.M., Lee M.K., Szabo C.I., Jerome N., McEuen M., Taylor M.,
RA Hood L., King M.-C.;
RT "Complete genomic sequence and analysis of 117 kb of human DNA
RT containing the gene BRCA1".
RL Genome Res. 6:1029-1049(1996).
RN [3]
RN FUNCTION AS A E2-DEPENDENT UBIQUITIN LIGASE.
RP MEDLINE=99432238; PubMed=10500182;
RX Lorick K.L., Jensen J.P., Fang S., Ong A.M., Hatakeyama S.,
RA Weissman A.M.;
RT "RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent
RT ubiquitination".
RL Proc. Natl. Acad. Sci. U.S.A. 96:11364-11369(1999).
RN [4]
RN IDENTIFICATION OF BRCA1 AS MEMBER OF BASC.
RP MEDLINE=20245492; PubMed=10783165;
RX Wang Y., Cortez D., Yazdi P., Neff N., Ellledge S.J., Qin J.;
RA "BASC, a super complex of BRCA1-associated proteins involved in the
RT recognition and repair of aberrant DNA structures".
RL Genes Dev. 14:927-939(2000).
RN [5]
RN INTERACTION WITH SMC1L1.
RP MEDLINE=21866464; PubMed=11877377;
RX Yazdi P.T., Wang Y., Zhao S., Patel N., Lee E.Y.-H.P., Qin J.;
RA "SMC1 is a downstream effector in the ATM/ATR1 branch of the human
RT S-phase checkpoint".
RL Genes Dev. 15:571-582(2002).
RN [6]
RN INTERACTION WITH NLF3.
RP MEDLINE=21601803; PubMed=11739404;
RX Ye Q., Hu Y.-F., Zhong H., Nye A.C., Belmont A.S., Li R.;
RA "BRCA1-induced large-scale chromatin unfolding and allele-specific
RT effects of cancer-predisposing mutations".
RL J. Cell Biol. 155:911-921(2001).
RN [7]
RN PHOSPHORYLATION BY ATM, AND MUTAGENESIS OF SER-1387; SER-1423 AND
RP SER-1524.
RX Xu B., O'Donnell A.H., Kim S.-T., Kastan M.B.;
RA "Phosphorylation of serine 1387 in BRCA1 is specifically required for
RT the Atm-mediated S-phase checkpoint after ionizing irradiation".
RL Cancer Res. 62:4588-4591(2002).
RN [8]
RN REVIEW ON VARIANTS.
RP MEDLINE=96400954; PubMed=8807330;
RX Couch F.J., Weber B.L.;
RA "Mutations and polymorphisms in the familial early-onset breast
RT cancer (BRCA1) gene".
RL Hum. Mutat. 8:8-18(1996).
RN [9]
RN VARIANTS LEU-1637; GLU-1708 AND ARG-1775.
RP MEDLINE=95025878; PubMed=7939630;
RX Futreal P.A., Liu Q., Shattuck-Eidens D., Cochran C., Harshman K.,
RA Tavtigian S., Bennett L.M., Haugen-Strano A., Swensen J., Miki Y.,
RA Edgington K., McClure M., Frye C., Weaver-Felhaus J., Ding W.,
RA Gholami Z., Soedarkrist P., Terry L., Jhanwar S., Berchuk A.,
RA Iglehart J.D., Marks J., Ballinger D.G., Barrett J.C., Skolnick M.H.,
RA Kamb A., Wiseman R.;
RT "BRCA1 mutations in primary breast and ovarian carcinomas".
RL Science 266:120-122(1994).
RN [10]
RP VARIANT BC GLY-64, AND VARIANTS ALA-772; ASN-1040 AND GLY-1443.
RX MEDLINE=95201806; PubMed=7894491;
RA Castilla L.H., Couch F.J., Erdos M.R., Hoskins K.F., Calzone K.,
RA Garber J.E., Boyd J., Lubin M.B., Deshano M.L., Brody L.C.,
RA Collins F.S., Weber B.L.;
RT "Mutations in the BRCA1 gene in families with early-onset breast and
RT ovarian cancer".
RL Nat. Genet. 8:387-391(1994).
RN [11]
RN VARIANT BC GLY-61, AND VARIANTS ARG-356; GLY-1038; ASN-1040; ARG-1183
RP AND GLY-1613.
RX MEDLINE=95201808; PubMed=7894493;
RA Friedman L.S., Oettermeyer E.A., Szabo C.I., Dowd P., Lynch E.D.,
RA Rowell S.E., King M.-C.;
RT "Confirmation of BRCA1 by analysis of germline mutations linked to
RT breast and ovarian cancer in ten families".
RL Nat. Genet. 8:399-404(1994).
RN [12]
RN VARIANT BC GLY-61.
RP MEDLINE=96108965; PubMed=8554067;
RX Serova O., Montagna M., Torchard D., Narod S.A., Tonin P., Sylla B.,
RA Lynch H.T., Feunteun J., Lenoir G.M.;
RT "A high incidence of BRCA1 mutations in 20 breast-ovarian cancer
RT families".
RL Am. J. Hum. Genet. 58:42-51(1996).
RN [13]
RN VARIANT BOC TRP-841.
RP MEDLINE=97123469; PubMed=8968716;
RX Barker D.F., Almeida E.F.A., Casey G., Fain P.R., Liao S.-Y.,
RA Masunaka I., Noble B., Kurosaki T., Anton-Culver H.;
RT "BRCA1 R841W: a strong candidate for a common mutation with moderate
RT phenotype".
RL Genet. Epidemiol. 13:595-604(1996).
RN [14]
RN VARIANTS BC AND BOC.
RP MEDLINE=96372821; PubMed=8776500;
RX Durocher F., Shattuck-Eidens D., McClure M., Labrie F.,
RA Skolnick M.H., Goldgar D.E., Simard J.;
RT "Comparison of BRCA1 polymorphisms, rare sequence variants and/or
RT missense mutations in unaffected and breast/ovarian cancer
RT populations".
RL Hum. Mol. Genet. 5:835-842(1996).
RN [15]
RN VARIANTS BC MET-271 AND SER-1150.
RX MEDLINE=96303704; PubMed=8723683;
RA Katagiri T., Eto I., Kobayashi K., Yoshimoto M., Iwase T.,
RA Kasumi F., Miki Y., Skolnick M.H., Nakamura Y.;
RT "Mutations in the BRCA1 gene in Japanese breast cancer patients".
RL Hum. Mutat. 7:334-339(1996).
RN [16]
RN VARIANT BC GLY-61, AND VARIANTS ARG-239; TRP-841 AND ILE-1512.
RX MEDLINE=98430998; PubMed=9760198;
RA Dong J., Chang-Claude J., Wu Y., Schumacher V., Debatin I., Tonin P.,
RA Royer-Pokora B.;
RT "A high proportion of mutations in the BRCA1 gene in German
RT breast/ovarian cancer families with clustering of mutations in the 3'
RT third of the gene".
RL Hum. Genet. 103:154-161(1998).
RN [17]
RN VARIANT BC GLY-64, AND VARIANTS ALA-772; GLU-820; ASN-1040; GLY-1443;
RP ILE-1512; LEU-1637 AND ILE-1652.
RX MEDLINE=98141685; PubMed=9482581;
RA Andersen T.I., Eiken H.G., Couch F., Kaada G., Skrede M., Johnsen H.,
RA Aloysius T.A., Tveit K.M., Tranbjærg L., Doerum A., Moeller P.,
RA Weber B.L., Boerresen-Dale A.-L.;
RT "Constant denaturation gel electrophoresis (CDGE) in BRCA1 mutation
RT screening".
RL Hum. Mutat. 11:166-174(1998).
RN [18]
RN VARIANTS BC SER-23; LEU-461; ASP-465; VAL-552; SER-892; ASP-960;
RP ILE-1025 AND ALA-1047.
RX MEDLINE=9827917; PubMed=9609997;
RA Katagiri T., Kasumi F., Yoshimoto M., Nomizu T., Asaishi K., Abe R.,

RA Tsuchiya A., Sugano M., Takai S., Yoneda M., Fukutomi T., Nanba K.,
RA Makita M., Okazaki H., Hirata K., Okazaki M., Furutsuma Y.,
RA Moriguchi Y., Iino Y., Karino T., Ayabe H., Hara S., Kajiwara T.,
RA Housa S., Shimizu T., Toda M., Yamazaki Y., Uchida T., Kunitomo K.,
RA Sano H., Kurebayashi J.-I., Shimotsu K., Nakamura Y., Waki Y.,
RT "High proportion of missense mutations of the BRCA1 and BRCA2 genes in
RL Japanese breast cancer families.";
RL J. Hum. Genet. 43:42-48(1998).
RN [19]
RN VARIANT OC ARG-1749.
RP MEDLINE=20455732; PubMed=10486320;
RA Gayther S.A., Russell P., Harrington P., Antoniou A.C., Easton D.F.,
RA Ponder B.A.J.,
RT "The contribution of germline BRCA1 and BRCA2 mutations to familial
RT ovarian cancer: no evidence for other ovarian cancer-susceptibility
RT genes.";
RL Am. J. Hum. Genet. 65:1021-1029(1999).
RN [20]
RN VARIANT BC SER-346, AND VARIANTS LEU-871; GLY-1038; ARG-1183 AND
RP GLY-1613
RX MEDLINE=99254821; PubMed=10323242;
RA Li S.-L., Tseng H.-M., Yang T.-P., Liu C.-H., Teng S.-J.,
RA Huang H.-W., Chen L.-M., Kao H.-W., Chen J.H., Tseng J.-N., Chen A.,
RA Hou M.-F., Huang T.-J., Chang H.-T., Mok K.-T., Tsai J.-H.,
RT "Molecular characterization of germline mutations in the BRCA1 and
RT BRCA2 genes from breast cancer families in Taiwan.";
RL Hum. Genet. 104:201-204(1999).
RN [21]
RN VARIANTS OC, AND VARIANTS.
RP MEDLINE=99214030; PubMed=10196379;
RA Janzic S.A., Ziogas A., Krutroy L.M., Krasner M., Plummer S.J.,
RA Cohen P., Gildea M., Barker D., Haile R., Casey G., Anton-Culver H.,
RT "Germline BRCA1 alterations in a population-based series of ovarian
RT cancer cases".
RL Hum. Mol. Genet. 8:889-897(1999).
CC -!- FUNCTION: Plays a central role in DNA repair by facilitating
CC cellular response to DNA repair. Required for appropriate cell
CC cycle arrests after ionizing irradiation in both the S-phase and

Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
-!- FUNCTION: Plays a central role in DNA repair by facilitating
cellular response to DNA repair. Required for appropriate cell
cycle arrests after ionizing irradiation in both the S-phase and
the G2 phase of the cell cycle. Involved in transcriptional
regulation of P21 in response to DNA damage. May function as a
transcriptional regulator. Mediates E2-dependent ubiquitination (By
similarity).
-!- SUBUNIT: Part of the BRCA1-associated genome surveillance complex
(BASC), which contains BRCA1, MSH2, MSH6, MLH1, ATM, BLM, PMS2 and
the RAD50-MRE11-NBS1 protein complex. This association could be a
dynamic process changing throughout the cell cycle and within
nuclear domains. CtIP interacts specifically with the BRCT
domains. Interacts with RNA polymerase II holoenzyme. Interacts
with SWI1 and COBRAL/NELEP (By similarity).
-!- SUBCELLULAR LOCATION: Nuclear (By similarity).
-!- PTM: Phosphorylated by ATM upon ionizing radiation (By
similarity).
-!- SIMILARITY: Contains 1 RING-type zinc finger.
-!- SIMILARITY: Contains 2 BRCT domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

EMBL; AF207822; AAG3492.1; -;
DR EMBL; AY365046; AAR04849.1; -;
DR GO; GO:0008274; C:gamma-tubulin ring complex; NAS.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0016329; F:apoptosis regulator activity; TAS.
DR GO; GO:0005515; F:protein binding; IPT.
DR GO; GO:0018563; F:transcriptional activator activity; TAS.
DR GO; GO:0016331; F:tubulin binding; NAS.
DR GO; GO:0006978; P:DNA damage response; signal transduction by . . ; TAS.
DR GO; GO:0046600; P:negative regulation of centriole replication; NAS.
DR GO; GO:0045739; P:positive regulation of DNA repair; NAS.
DR GO; GO:0042127; P:regulation of cell proliferation; TAS.
DR GO; GO:0006359; P:regulation of transcription from Pol III pr. . . ; TAS.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR002378; Brst_cancerI.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00533; BRCT; 2.
DR PRINTS; PR00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00292; BRCT; 2.
DR PROSITE; PS00172; BRCT; 2.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR DNA repair; Zinc-finger; DNA-binding; Nuclear protein; Anti-oncogene;
KW Repeat; Phosphorylation; Polymorphism.
FT ZN_FING 24 65
FT DOMAIN 1642 1736
FT DOMAIN 1756 1855
FT DOMAIN 651 654
FT MOD_RES 1387 1387
FT MOD_RES 1423 1423
FT MOD_RES 1524 1524
FT VARIANT 309 309
FT VARIANT 590 590
FT VARIANT 731 731
FT VARIANT 1100 1100
FT CONFLICT 427 427
FT CONFLICT 925 925
FT CONFLICT 1520 1520
SQ SEQUENCE 1863 AA; 207898 MW; 49673829CCFA756E CRC64;

Query Match 81.6%; Score 31; DB 1; Length 1863;

Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 AADPHE 7
|:|||||
Db 433 ASDPHE 438

RESULT 15

Y126_MYCTU STANDARD; PRT; 232 AA.
AC Q50702;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical PPE-family protein RV3426.
GN RV3426 OR MYCY78.03C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=9829597; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham D., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z77165; CAB01030.1; -.
DR PIR; G70738; G70738.
DR TuberculList; RV3426; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 232 AA; 25872 MW; D76512D49EB272C6 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 232;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AADPHE 7
|:|||||
Db 169 AADPQE 175

Search completed: April 19, 2004, 11:52:54
Job time : 2.51062 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 2.70175 Seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_362_368

Perfect score: 38

Sequence: 1 AAADPHE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	396	Q81UK7	Q81uk7 homo sapien
2	38	100.0	417	Q86YGO	Q86ygo homo sapien
3	38	100.0	4177	Q9GUF2	Q9guf2 caenorhabdi
4	34	89.5	157	Q8NF55	Q8nf55 homo sapien
5	34	89.5	310	Q7Z3Y1	Q7z3y1 homo sapien
6	34	89.5	323	Q68908	Q68908 streptomyc
7	34	89.5	323	Q8X200	Q8x200 ralstonia s
8	34	89.5	691	Q7ZUR6	Q7zue6 brachydanio
9	34	89.5	740	Q28905	Q28905 archaeoglob
10	33	86.8	180	Q8PRI4	Q8pr14 xanthomonas
11	33	86.8	238	Q8RUD4	Q8rud4 deinooccus
12	33	86.8	283	Q8PC99	Q8pc99 xanthomonas
13	33	86.8	308	Q9LQ92	Q9lg92 arabidopsis
14	33	86.8	351	Q8U7L3	Q8u7l3 agrobacteri
15	33	86.8	414	Q53279	Q53279 mycobacteri
16	33	86.8	414	Q7TXE8	Q7txe8 mycobacteri

17	33	86.8	438	16	Q9CBR3	Q9cbr3 mycobacteri
18	33	86.8	459	10	Q8WZB7	Q8wzb7 zea mays (m
19	33	86.8	464	10	Q8WZB6	Q8wzb6 zea mays (m
20	33	86.8	630	16	Q89X41	Q89x41 bradyrhizob
21	33	86.8	632	2	Q8KWS3	Q8kws3 pectobacter
22	33	86.8	667	2	Q9WXA5	Q9wxa5 erwinia car
23	33	86.8	749	11	Q8C3X9	Q8c3x9 mus musculu
24	33	86.8	768	5	Q97215	Q97215 leishmania
25	33	86.8	1205	11	Q8CGW1	Q8cgw1 mus musculu
26	33	86.8	1574	11	Q9QXC2	Q9qxc2 mus musculu
27	33	86.8	1627	10	Q8LQW0	Q8lq10 oryza sativ
28	33	86.8	1939	2	Q8GME1	Q8gme1 streptomyc
29	32	84.2	130	2	Q9AES9	Q9aes9 aeromonas h
30	32	84.2	130	2	Q8GLR0	Q8glr0 aeromonas s
31	32	84.2	359	16	Q8R642	Q8r642 fusobacteri
32	32	84.2	363	17	Q28245	Q28245 archaeoglob
33	32	84.2	421	16	Q8U0Z0	Q8u0z0 agrobacteri
34	31	81.6	101	16	Q8DK11	Q8dk11 synechococc
35	31	81.6	105	17	Q9YBU3	Q9ybu3 aeropyrum p
36	31	81.6	248	16	Q88T94	Q88t94 lactobacill
37	31	81.6	272	5	Q45091	Q45091 caenorhabdi
38	31	81.6	325	16	Q82BG7	Q82bg7 streptomyc
39	31	81.6	345	5	P91437	P91437 caenorhabdi
40	31	81.6	390	13	Q42223	Q42223 trachemys s
41	31	81.6	418	10	Q9SGS7	Q9sgs7 arabidopsis
42	31	81.6	477	4	Q8N9F1	Q8n9f1 homo sapien
43	31	81.6	493	16	Q7V4S5	Q7v4s5 prochloroco
44	31	81.6	499	10	Q9LG10	Q9lg10 arabidopsis
45	31	81.6	630	13	Q57336	Q57336 petromyzon

ALIGNMENTS

RESULT 1

Q81UK7 PRELIMINARY; PRT; 396 AA.

AC Q81UK7;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R;
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC035969; AAH35969.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport P10T; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 38; DB 4; Length 396;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7

Db 173 AAADPHE 179

RESULT 2

Q86YGO

```

ID Q86YGO PRELIMINARY; PRT; 417 AA.
AC Q86YGO;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47360 MW; 16E76483BEF4E8D CRC64;

Query Match 100.0%; Score 38; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
DB 194 AAADPHE 200

RESULT 3
Q86YGO PRELIMINARY; PRT; 417 AA.
ID Q86YGO;
AC Q86YGO;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein Y67D8C.5.
GN Y67D8C.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodoridae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA Waterston R.;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Edwards J., Du H., Lamar B., Kemp K., Wohldmann P., Walker C.;
RT "The sequence of C. elegans cosmid Y67D8C.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC025724; BAG23375.2; -.
DR WormPep; Y67D8C.5; C33165.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.

```

```

DR InterPro; IPR008938; ARM.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00627; UBA; 1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS0030; UBA; 1.
KW Hypothetical protein.
SQ SEQUENCE 4177 AA; 465401 MW; 9C32EF90AB87FD58 CRC64;

Query Match 100.0%; Score 38; DB 5; Length 4177;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
DB 1431 AAADPHE 1437

RESULT 4
Q8NF55 PRELIMINARY; PRT; 157 AA.
ID Q8NF55;
AC Q8NF55;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE FLJ00335 protein (Fragment).
GN FLJ00335.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
spleen.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK090424; BAC03405.1; -.
DR InterPro; IPR000949; ELM2.
DR Pfam; PF01448; ELM2; 1.
FT NON_TER 1
SQ SEQUENCE 157 AA; 17936 MW; 8DFE8C792B17D1BE CRC64;

Query Match 89.5%; Score 34; DB 4; Length 157;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
DB 32 AAADPHK 38

RESULT 5
Q723Y1 PRELIMINARY; PRT; 310 AA.
ID Q723Y1;
AC Q723Y1;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

```

Matches	6;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

QY	1	AAADPHE 7							
Db	49	AAVDPHE 55							

RESULT 7									
Q8XZ00		PRELIMINARY;		PRT;		323	AA.		
ID	Q8XZ00								
AC	Q8XZ00;								
DT	01-MAR-2002 (TrEMBLrel. 20, Created)								
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)								
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)								
DE	Hypothetical protein RSC1606.								
GN	RSC1606 OR RS03969.								
OS	Ralstonia solanacearum (Pseudomonas solanacearum).								
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;								
OC	Burkholderiaceae; Ralstonia.								
OX	NCBI_TaxID=305;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=GM11000;								
RC	MEDLINE=21681879; PubMed=11823852;								
RA	Salanoubat M., Genin S., Ariguenave F., Gouzy J., Mangenot S.,								
RA	Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,								
RA	Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,								
RA	Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,								
RA	Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,								
RA	Weissenbach J., Boucher C.A.;								
RT	"Genome sequence of the plant pathogen Ralstonia solanacearum.";								
RL	Nature 415:497-502(2002).								
DR	EMBL; AL646065; CAD15308.1; "								
DR	InterPro; IPR000286; His deacetylase.								
DR	Pfam; PF00850; Hist deacetyl; 1.								
DR	PRINTS; PR01270; HDASUPER.								
KW	Hypothetical protein; Complete proteome.								
SK	SEQUENCE 323 AA; 34637 MW; 304513B9155416DB CRC64;								

Query Match	89.5%;	Score 34;	DB 16;	Length 323;		
Best Local Similarity	85.7%;	Pred. No. 73;				
Matches	6;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	AAADPHE 7							
Db	234	AGADPHE 240							

RESULT 8									
Q7ZUE6		PRELIMINARY;		PRT;		691	AA.		
ID	Q7ZUE6								
AC	Q7ZUE6;								
DT	01-JUN-2003 (TrEMBLrel. 24, Created)								
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)								
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)								
DE	Similar to guanine nucleotide-releasing factor 2 (Specific for crk proto-oncogene).								
DE	Brachydanio rerio (Zebrafish) (Danio rerio).								
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;								
OC	Cyprinidae; Danio.								
OX	NCBI_TaxID=7955;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Body;								
RA	Strausberg R.;								
RL	Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; BC049305; AAH49305.1; "								
SK	SEQUENCE 691 AA; 76263 MW; 90CC6764E6D89EE7 CRC64;								

Query Match	89.5%;	Score 34;	DB 13;	Length 691;		
Best Local Similarity	100.0%;	Pred. No. 1.6e+02;				
Matches	6;	Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
OY 2 AADPHE 7
Db 560 AADPHE 565

RESULT 9
O28905 PRELIMINARY; PRT; 740 AA.
AC O28905;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hydrogenase expression/formation regulatory protein (HYPF).
GN AF1366.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49559;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervilange A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
DR EMBL; AE001009; AAB89876.1; -.
DR TIGR; AF1366. -.
DR GO; GO:0003988; F:acylphosphatase activity; IEA.
DR InterPro; IPR001792; Acylphosphatase.
DR InterPro; IPR004421; HyPF.
DR InterPro; IPR006071; SUA5/YciO/YrdC.
DR InterPro; IPR006070; SUA5/YciO/YrdC_N.
DR Pfam; PF00708; Acylphosphatase; 1.
DR Pfam; PF01300; SUA5_YciO_YrdC; 1.
DR ProDom; PD001884; Acylphosphatase; 1.
DR ProDom; PD002209; SUA5/YciO/YrdC; 1.
DR TIGRFAMs; TIGR00143; hyPF; 1.
DR PROSITE; PS00150; ACYLPHOSPHATASE_1; 1.
DR PIRSF; PIRSF002256; HyPF; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 740 AA; 83060 MW; 9C0D5320A64B157 CRC64;

Query Match 89.5%; Score 34; DB 17; Length 740;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 AADPHE 7
Db 606 AADPHE 612

RESULT 10
Q8PRL4 PRELIMINARY; PRT; 180 AA.
AC Q8PRL4;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein XACB0006.
GN XACB0006.
OS Xanthomonas axonopodis (pv. citri).
```

```
OY Plasmid pXAC64.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OX Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Cimarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katayama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Localsi E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.H.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C. de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RA "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AE008925; RAM39252.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 180 AA; 20041 MW; C317CAEEA3C31E86 CRC64;

Query Match 86.8%; Score 33; DB 16; Length 180;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AADPHE 7
Db 102 AADPHE 108

RESULT 11
Q9RUD4 PRELIMINARY; PRT; 238 AA.
AC Q9RUD4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein DR1454.
GN DR1454.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OX Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Bisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RA "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL; AE001990; AAP11027.1; -.
DR PIR; H75392; H75392.
DR TIGR; DR1454; -.
DR InterPro; IPR008941; TPR-like.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 238 AA; 25113 MW; 489BBC792C11E7AD CRC64;
```

```

Query Match      86.8%; Score 33; DB 16; Length 238;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPH 6
      |||||
Db      162 AAADPH 167

RESULT 12
Q8PC99 PRELIMINARY; PRT; 283 AA.
AC Q8PC99
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Proteorhizobium oxidase.
GN HEMK OR XCC0835
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
CX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavini F., Cardoso J., Chamberg F., Ciapina L.P.,
RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Fornighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AE012183; AAM40150.1; -
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0008276; F:protein methyltransferase activity; IEA.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR GO; GO:0006479; P:protein amino acid methylation; IEA.
DR InterPro; IPR004556; HemK.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR000051; SAW_bind.
DR TIGRFAMs; TIGR000336; hemKfam; 1.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Complete proteome.
SQ SEQUENCE 283 AA; 30275 MW; A5F4E951E8E33479 CRC64;

Query Match      86.8%; Score 33; DB 16; Length 283;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPH 6
      |||||
Db      191 AAADPH 196

RESULT 13
Q9LQ92 PRELIMINARY; PRT; 308 AA.
AC Q9LQ92
ID Q9LQ92
Q9LQ92;

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Tln6.5 protein.
GN Tln6.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=sv. Columbia;
RA Liu S., Chan A., Yu G., Lee J., Lenz C., Pham P., Sakano H.,
RA Toriumi M., Chin C., Chiou J., Choi E., Chung M., Gonzalez E.,
RA Huang B., Liu A., Vaysberg M., Altati H., Brooks S., Buehler E.,
RA Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C., Shinn P.,
RA Southwick A., Davis R., Ecker J., Federspiel N., Theologis A.;
RT "The sequence of BAC Tln6 from Arabidopsis thaliana chromosome 1."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009273; AAF78398.1; -
DR PIR; P86147; F86147.
DR InterPro; IPR003613; Znf_modRING.
DR Pfam; PF04564; U-box; 1.
DR SMART; SM00504; Ubox; 1.
SQ SEQUENCE 308 AA; 35392 MW; CE53CFA27C77547D CRC64;

Query Match      86.8%; Score 33; DB 10; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPH 6
      |||||
Db      117 AAADPH 122

RESULT 14
Q8U7L3 PRELIMINARY; PRT; 351 AA.
AC Q8U7L3
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE IS5 family transposase.
GN TNP OR ATU4436 OR AGR_L_863GL.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P., Woo L.,
RA Okura V.K., Zhou Y., Chen L., Wood G.B., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RN Science 294:2317-2323(2001).
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Hummel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,

```

RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.,
RT Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL; AE009372; AAL45230.1; ALT_INIT.
DR EMBL; AE008242; AAK89004.1; -.
DR PIR; AH3101; AH3101.
DR PIR; B98185; B98185.
DR InterPro; IPR002559; Transposase_11.
DR Pfam; PF01609; Transposase_11; 1.
KW Complete proteome.
SQ SEQUENCE 351 AA; 39162 MW; 8C580BA75CBDF249 CRC64;

Query Match 86.8%; Score 33; DB 16; Length 351;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAADPHE 7
| | | | |
Db 173 ALADPHE 179

RESULT 15

OS3279 PRELIMINARY; FRT; 414 AA.
AC OS3279;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Glycosyl transferase).
GN RV3032 OR MT012.47 OR MT3116.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV.
RX MEDLINE=9825987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AL021287; CAAL6117.1; -.
DR EMBL; AE007130; AAK47446.1; -.
DR PIR; C70859; C70859.
DR TIGR; MT3116; -.
DR TubercuList; Rv3032; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biogenesis; IEA.
DR InterPro; IPR001296; Glyco.trans.1.
DR Pfam; PF00534; Glycos.transf.1; 1.
KW Hypothetical protein; Transferase; Complete proteome.
SQ SEQUENCE 414 AA; 44805 MW; 64ADC7A4EA17A0DE CRC64;

Query Match 86.8%; Score 33; DB 16; Length 414;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AAADPHE 7
| | | | |
Db 66 AAQDPHE 72

Search completed: April 19, 2004, 12:00:11
Job time : 4.70175 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 4.11727 Seconds
(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_362_368

Perfect score: 38

Sequence: 1 AAADPHE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	38	100.0	13	4	ABBS6230 Vascular
2	38	100.0	13	4	ABBS6225 Vascular
3	38	100.0	13	4	AAU25032 Schizophr
4	38	100.0	13	4	AAU15376 Schizophr
5	38	100.0	103	4	AAU33145 Novel hum
6	38	100.0	112	4	AAO2636 Human pol
7	38	100.0	188	3	AAU33145 Novel hum
8	38	100.0	228	3	AAU33145 Novel hum
9	38	100.0	243	4	AAU33087 Novel hum
10	38	100.0	243	4	AAU33081 Novel hum
11	38	100.0	293	4	AAU33082 Novel hum
12	38	100.0	373	1	AAU33087 N-termina
13	38	100.0	388	1	AAU33087 N-termina
14	38	100.0	389	1	AAU33087 N-termina
15	38	100.0	390	1	AAU33087 N-termina
16	38	100.0	407	1	AAU33087 N-termina
17	38	100.0	463	2	AAU33087 N-termina
18	38	100.0	507	4	AAU33286 Novel hum
19	38	100.0	507	4	AAU33294 Novel hum
20	38	100.0	550	4	AAU33294 Novel hum
21	38	100.0	584	6	ABG72381 Mature hu
22	38	100.0	585	1	ABG72381 Sequence
23	38	100.0	585	1	ABG72381 Mature hu
24	38	100.0	585	1	ABG72381 Mature hu
25	38	100.0	585	2	AAU5318 Human ser

26	38	100.0	585	2	AAU5318 Human ser
27	38	100.0	585	2	AAU5318 Human ser
28	38	100.0	585	2	AAU5318 Human ser
29	38	100.0	585	2	AAU5318 Human ser
30	38	100.0	585	2	AAU5318 Human ser
31	38	100.0	585	2	AAU5318 Human ser
32	38	100.0	585	2	AAU5318 Human ser
33	38	100.0	585	2	AAU5318 Human ser
34	38	100.0	585	2	AAU5318 Human ser
35	38	100.0	585	2	AAU5318 Human ser
36	38	100.0	585	2	AAU5318 Human ser
37	38	100.0	585	2	AAU5318 Human ser
38	38	100.0	585	2	AAU5318 Human ser
39	38	100.0	585	2	AAU5318 Human ser
40	38	100.0	585	2	AAU5318 Human ser
41	38	100.0	585	2	AAU5318 Human ser
42	38	100.0	585	2	AAU5318 Human ser
43	38	100.0	585	2	AAU5318 Human ser
44	38	100.0	585	2	AAU5318 Human ser
45	38	100.0	585	2	AAU5318 Human ser

ALIGNMENTS

RESULT 1
ABBS6230
ID ABBS6230 standard; peptide; 13 AA.
XX
AC ABBS6230;
XX
DT 15-FEB-2002 (first entry)
XX
DE Vascular dementia-associated protein isoform (VPI) 430.
XX
KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
XX
KW diagnosis; prognosis; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200169261-A2.
XX
PD 20-SEP-2001.
XX
PF 14-MAR-2001; 2001WO-GB001106.
XX
PR 15-MAR-2000; 2000GB-00006285.
PR 24-NOV-2000; 2000GB-00028734.
PR 28-NOV-2000; 2000US-00724391.
XX
(OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
Herath HMCAC, Parekh RB, Rohlf C;
WPI; 2001-557937/62.

Screening, diagnosis or prognosis of vascular dementia (VD), useful for determining stage of VD and monitoring the effect of VD therapy, for comprises analyzing body fluid by 2-dimensional electrophoresis for features correlated with VD.
Claim 6; Page 39; 151pp; English.
The invention relates to screening, diagnosis or prognosis of Vascular Dementia (VD) in a subject comprising analysing body fluid from the subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of features containing at least one chosen feature whose relative abundance correlates with the presence, absence, stage or severity of VD or predicts the onset or course of VD, especially detecting in a sample of cerebrospinal fluid (CSF) from the subject one of 223 VD-associated protein isoforms (VPIs) (ABBS5801-ABBS6295) as fully defined in the specification. Detecting VD-associated features and VPI is useful for screening, diagnosis or prognosis of VD, for determining the stage or

CC severity of VD, for identifying a subject at risk of VD or for monitoring
 CC the effect of therapy administered to a subject having VD. Nucleic acids
 CC encoding a VPI or inhibiting the function of a VPI are useful for the
 CC treatment of VD and for gene therapy
 XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 38; DB 4; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAADPHE 7
 |||||
 Db 3 AAADPHE 9

RESULT 2

ABBS6225
 ID ABB56225 standard; peptide; 13 AA.

XX
 AC ABB56225;

XX
 DT 15-FEB-2002 (first entry)

XX
 DE Vascular dementia-associated protein isoform (VPI) 425.

XX
 KW Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
 KW diagnosis; prognosis; gene therapy.

XX
 OS Homo sapiens.

XX
 PN WO200169261-A2.

XX
 PD 20-SEP-2001.

XX
 PF 14-MAR-2001; 2001WO-GB001106.

XX
 PR 15-MAR-2000; 2000GB-00006285.

XX
 PR 24-NOV-2000; 2000GB-00028734.

XX
 PR 28-NOV-2000; 2000US-00724391.

XX
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX
 PI Herath HMAc, Parekh RB, Rohlf C;

XX
 DR WPI; 2001-557937/62.

XX
 PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
 PT determining stage of VD and monitoring the effect of VD therapy,
 PT comprises analyzing body fluid by 2-dimensional electrophoresis for
 PT features correlated with VD.

XX
 PS Claim 6; Page 39; 151pp; English.

XX
 CC The invention relates to screening, diagnosis or prognosis of Vascular
 CC Dementia (VD) in a subject comprising analysing body fluid from the
 CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
 CC features containing at least one chosen feature whose relative abundance
 CC correlates with the presence, absence, stage or severity of VD or
 CC predicts the onset or course of VD, especially detecting in a sample of
 CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
 CC protein isoforms (VPIs) (ABB5801-ABB56295) as fully defined in the
 CC specification. Detecting VD-associated features and VPI is useful for the
 CC screening, diagnosis or prognosis of VD, for determining the stage or
 CC severity of VD, for identifying a subject at risk of VD or for monitoring
 CC the effect of therapy administered to a subject having VD. Nucleic acids
 CC encoding a VPI or inhibiting the function of a VPI are useful for the
 CC treatment of VD and for gene therapy

XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 38; DB 4; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.57;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AAADPHE 7
 |||||
 Db 3 AAADPHE 9

RESULT 3

AAU25032
 ID AAU25032 standard; peptide; 13 AA.

XX
 AC AAU25032;

XX
 DT 18-DEC-2001 (first entry)

XX
 DE Schizophrenia-Associated Protein Isoform (SPI) peptide #261.

XX
 KW Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-238; SPI-240;
 KW neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.

XX
 OS Homo sapiens.

XX
 PN WO200162785-A2.

XX
 PD 30-AUG-2001.

XX
 PF 23-FEB-2001; 2001WO-GB000792.

XX
 PR 24-FEB-2000; 2000GB-00004415.

XX
 PR 28-DEC-2000; 2000US-00750395.

XX
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX
 PI Herath HMAc, Parekh RB, Rohlf C, Terrett JA, Tyson KL;

XX
 DR WPI; 2001-570624/64.

XX
 PT New schizophrenia associated protein isoforms and encoding nucleic acid
 PT molecules, useful for treatment, diagnosis and prognosis of schizophrenia
 PT and screening for potential drugs for treatment and new drug targets.

XX
 PS Disclosure; Page 34; 148pp; English.

XX
 CC The sequence represents a schizophrenia-associated protein isoform (SPI).
 CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable
 CC in cerebrospinal fluid, serum or plasma and are useful markers of
 CC schizophrenia. The sequences can be used for treatment and diagnosis of
 CC schizophrenia. Screening, prognosis, monitoring the results of therapy,
 CC identifying patients most likely to respond to a particular therapy and
 CC identification of new targets for drug treatment. SPI DNA is useful as a
 CC nucleic acid probe to detect the presence of nucleic acids or SPIs

XX
 SQ Sequence 13 AA;

Query Match 100.0%; Score 38; DB 4; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAADPHE 7
 |||||
 Db 3 AAADPHE 9

RESULT 4

AAU15376

ID AAU15376 standard; peptide; 13 AA.

XX
 AC AAU15376;

XX
 DT 24-OCT-2001 (first entry)

XX
 DE Schizophrenia-associated isoform peptide #261.

KW Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;
 KW neurological disorder; neuropathy.
 XX Homo sapiens.
 OS WO200163293-A2.
 PN 30-AUG-2001.
 XX 23-FEB-2001; 2001WO-GB000783.
 XX 24-FEB-2000; 2000GB-00004415.
 XX 28-DEC-2000; 2000US-00750395.
 XX (OXFO-) OXFORD GLYSCSCIENCES UK LTD.
 XX Herath HMAc, Parekh RB, Rohlf C;
 XX WPI; 2001-502868/55.
 XX Diagnosing and monitoring Schizophrenia by detecting the presence of
 PT Schizophrenia Associated Features and Schizophrenia Associated Protein
 PT isoforms in samples of cerebrospinal fluid.
 XX Claim 6; Page 34; 160pp; English.
 XX The invention relates to methods and compositions for screening,
 CC diagnosis and prognosis of Schizophrenia. The method involves detecting
 CC the presence of Schizophrenia (SCH) Associated Features (SfAs) and SCH
 CC Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,
 CC immunoblotting or hybridisation assay, for diagnosing and monitoring SCH,
 CC studying the effectiveness of treatments and for identifying potential
 CC therapeutic agents. The method is used for (1) screening or diagnosis of
 CC SCH and the relative abundance of at least 1 chosen feature correlates
 CC with the presence or absence of SCH; and (2) monitoring the effect of
 CC therapy administered to a subject with SCH and the relative abundance of
 CC at least 1 chosen feature which correlates with the severity of SCH. The
 CC expression and activity of the SfAs, SPIs and related molecules (e.g.
 CC secondary messengers) are studied to diagnose SCH, monitor the progress
 CC of the disorder and the effectiveness of treatment and as targets to
 CC identify and produce potential therapeutic agents for the treatment of
 CC SCH. The paucity of detectable neurologic defects distinguishes
 CC neuropsychiatric disorders such as SCH from neurological disorders, where
 CC manifestations of anatomical and biochemical changes have been identified
 CC in many cases. Consequently the identification and characterisation of
 CC cellular and/or molecular causative defects and neuropathies are
 CC necessary for improved treatment of neuropsychiatric disorders. AAU1514-
 CC AAU15142 represent the amino acid sequences of schizophrenia-associated
 CC isoforms used in the method of the invention
 XX SQ Sequence 13 AA;
 Query Match 100.0%; Score 38; DB 4; Length 13;
 Best Local Similarity 100.0%; Pred. No. 0.57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAADPHE 7
 |||||
 Db 3 AAADPHE 9
 RESULT 5
 AAU33145
 ID AAU33145 standard; protein; 103 AA.
 XX AAU33145;
 XX 18-DEC-2001 (first entry)
 XX Novel human secreted protein #3636.
 XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX Homo sapiens.
 OS WO200179449-A2.
 PN 25-OCT-2001.
 XX 16-APR-2001; 2001WO-US008656.
 XX 18-APR-2000; 2000US-00552929.
 XX 26-JAN-2001; 2001US-00770160.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-611725/70.
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX Claim 20; Page 712; 765pp; English.
 XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX SQ Sequence 103 AA;
 Query Match 100.0%; Score 38; DB 4; Length 103;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAADPHE 7
 |||||
 Db 34 AAADPHE 40
 RESULT 6
 AA02636
 ID AA02636 standard; protein; 112 AA.
 XX AA02636;
 XX 06-NOV-2001 (first entry)
 XX Human polypeptide SEQ ID NO 16528.
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX Homo sapiens.
 XX WO200164835-A2.
 XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.
 PF 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-514838/56.
 XX N-PSDB; AAI82567.
 DR Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 XX Claim 20; SEQ ID NO 16528; 1399pp + Sequence Listing; English.
 XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 112 AA;
 SQ

Query Match 100.0%; Score 38; DB 4; Length 112;
 Best Local Similarity 100.0%; Pred. No. 5.9; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
 DB 85 AAADPHE 91
 |||||
 |||||

RESULT 7
 AAY83948
 ID AAY83948 standard; protein; 188 AA.
 AC AAY83948;
 XX 28-JUL-2000 (first entry)
 DT Yeast codon-biased recombinant HSA protein fragment HSA-II.
 DE Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
 KW overlapping oligonucleotide; expression vector.
 XX Homo sapiens.
 OS Synthetic.
 XX CN1239103-A.
 PN 22-DEC-1999.
 XX 17-JUN-1998; 98CN-00102506.
 PF 17-JUN-1998; 98CN-00102506.
 XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.
 PA Li S, Lu D;
 PI WPI; 2000-351198/31.
 XX N-PSDB; AAI10094.
 DR Process for preparing recombinant human serum albumin comprising yeast
 PT biased sex codons - uses a recombinant DNA technique.
 XX Example 1; Fig 7; 44pp; Chinese.
 PS The method relates to a method of recombinantly producing human serum
 CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
 CC a yeast codon bias. The complete HSA gene (AAAI0091) was generated as
 CC three synthetic fragments (AAAI0092-AI0094) joined by recombinant DNA
 CC technology. Each HSA fragment was synthesised from overlapping
 CC oligonucleotide fragments that were extended. This sequence represents
 CC the sequence of the HSA fragment HSA-II encoded by the human gene with a
 CC yeast codon bias. The invention also covers a recombinant expression and
 CC vector, yeast host cells carrying the recombinant expression vector and
 CC the process for producing human serum albumin in the yeast host cell,
 CC especially in secretory mode
 XX

XX Process for preparing recombinant human serum albumin comprising yeast
 PT biased sex codons - uses a recombinant DNA technique.
 XX Example 1; Fig 5; 44pp; Chinese.
 PS The method relates to a method of recombinantly producing human serum
 CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
 CC three synthetic fragments (AAAI0092-AI0094) joined by recombinant DNA
 CC technology. Each HSA fragment was synthesised from overlapping
 CC oligonucleotide fragments that were extended. This sequence represents
 CC the sequence of the HSA fragment HSA-II encoded by the human gene with a
 CC yeast codon bias. The invention also covers a recombinant expression and
 CC vector, yeast host cells carrying the recombinant expression vector and
 CC the process for producing human serum albumin in the yeast host cell,
 CC especially in secretory mode
 XX

Query Match 100.0%; Score 38; DB 3; Length 188;
 Best Local Similarity 100.0%; Pred. No. 10; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
 DB 178 AAADPHE 184
 |||||
 |||||

RESULT 8
 AAY83949
 ID AAY83949 standard; protein; 228 AA.
 AC AAY83949;
 XX 28-JUL-2000 (first entry)
 DT Yeast codon-biased recombinant HSA protein fragment HSA-III.
 DE Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
 KW overlapping oligonucleotide; expression vector.
 XX Homo sapiens.
 OS Synthetic.
 XX CN1239103-A.
 PN 22-DEC-1999.
 XX 17-JUN-1998; 98CN-00102506.
 PF 17-JUN-1998; 98CN-00102506.
 XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.
 PA Li S, Lu D;
 PI WPI; 2000-351198/31.
 XX N-PSDB; AAI10094.
 DR Process for preparing recombinant human serum albumin comprising yeast
 PT biased sex codons - uses a recombinant DNA technique.
 XX Example 1; Fig 7; 44pp; Chinese.
 PS The method relates to a method of recombinantly producing human serum
 CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
 CC a yeast codon bias. The complete HSA gene (AAAI0091) was generated as
 CC three synthetic fragments (AAAI0092-AI0094) joined by recombinant DNA
 CC technology. Each HSA fragment was synthesised from overlapping
 CC oligonucleotide fragments that were extended. This sequence represents
 CC the sequence of the HSA fragment HSA-III encoded by the human gene with a
 CC yeast codon bias. The invention also covers a recombinant expression
 CC

CC vector, yeast host cells carrying the recombinant expression vector and
CC the process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
SQ Sequence 228 AA;

Query Match 100.0%; Score 38; DB 3; Length 228;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAADPHE 7
Db 5 AAADPHE 11

RESULT 9

AAU33087
ID AAU33087 standard; protein; 243 AA.

XX AC AAU33087;
XX 18-DEC-2001 (first entry)
XX Novel human secreted protein #3578.
XX Human; vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX WO200179449-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US008656.
XX 18-APR-2000; 2000US-00552929.
XX 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX Claim 20; Page 706; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX

SQ Sequence 243 AA;

Query Match 100.0%; Score 38; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAADPHE 7
Db 64 AAADPHE 70

RESULT 10

AAU33081
ID AAU33081 standard; protein; 245 AA.

XX AC AAU33081;
XX 18-DEC-2001 (first entry)
XX Novel human secreted protein #3572.
XX Human; vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX WO200179449-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US008656.
XX 18-APR-2000; 2000US-00552929.
XX 26-JAN-2001; 2001US-00770160.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-611725/70.
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX Claim 20; Page 705; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX

SQ Sequence 245 AA;

Query Match 100.0%; Score 38; DB 4; Length 245;
Best Local Similarity 100.0%; Pred. No. 14;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAADPHE 7
Db 15 AAADPHE 21

RESULT 11

```

XX DE N-terminal of human serum albumin polypeptide.
XX XX
XX KW Human serum albumin polypeptide; plasma expanders.
XX OS Homo sapiens; (Human).
XX PN EP322094-A.
XX PD 28-JUN-1989.
XX PF 25-OCT-1988; 88EP-00310000.
XX PR 30-OCT-1987; 87GB-00025529.
XX PA (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX DR WPI; 1989-186464/26.
XX XX
XX XX New N-terminal fragments of human serum albumin - esp. useful as blood
XX PT plasma expanders.
XX PS Claim 2; Page 9; 20pp; English.
XX CC N-terminal portion of human serum albumin. Used as plasma expanders, or
XX CC as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-
XX CC MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
XX CC field)
XX SQ Sequence 373 AA;

Query Match 100.0%; Score 38; DB 1; Length 373;
Best Local Similarity 100.0%; Pred.No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps

Qy 1 AAADPHE 7
Db 362 AAADPHE 368

RESULT 13
AAP90389
ID AAP90389 standard; protein; 388 AA.
XX AAP90389;
AC AC
XX AC
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX DT
XX DE N-terminal human serum albumin polypeptide.
XX DE N-terminal human serum albumin polypeptide; plasma expanders.
XX KW Homo sapiens; (Human).
XX OS Homo sapiens; (Human).
XX PN EP322094-A.
XX PD 28-JUN-1989.
XX PF 25-OCT-1988; 88EP-00310000.
XX PR 30-OCT-1987; 87GB-00025529.
XX PA (DELZ ) DELTA BIOTECHNOLOGY LTD.
XX PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX DR WPI; 1989-186464/26.
XX XX
XX XX New N-terminal fragments of human serum albumin - esp. useful as blood
XX PT plasma expanders.

```

```
PT plasma expanders.
XX
XX Claim 2; Page 9; 20pp; English.
PS
CC N-terminal fragment of human serum albumin used as plasma expander, or as
CC substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
XX Sequence 388 AA;
SQ
    Query Match          100.0%; Score 38; DB 1; Length 388;
    Best Local Similarity 100.0%; Pred. No. 22;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAADPHE 7
Db      362 AAADPHE 368
      |||||
RESULT 14
AAP90390
ID AAP90390 standard; protein; 389 AA.
XX
XX AAP90390;
AC
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
XX N-terminal human serum albumin.
DE
XX
XX N-terminal portion of human serum albumin; plasma expanders.
KW
XX
XX Homo sapiens; (Human).
OS
XX
XX EP322094-A.
PN
XX 28-JUN-1989.
PD
XX 25-OCT-1988; 88EP-00310000.
PF
XX 30-OCT-1987; 87GB-00025529.
PR
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
PI
XX WPI; 1989-186464/26.
DR
XX
XX New N-terminal fragments of human serum albumin - esp. useful as blood
PT plasma expanders.
FT
XX
XX Claim 2; Page 9; 20pp; English.
PS
XX N-terminal portion of human serum albumin. Used to make new N-terminal
XX fragments which are used as plasma expanders, or as substitutes for HSA
CC or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct PA
CC field.) (Updated on 24-OCT-2003 to standardise OS field)
CC
XX
XX Sequence 390 AA;
SQ
    Query Match          100.0%; Score 38; DB 1; Length 390;
    Best Local Similarity 100.0%; Pred. No. 23;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAADPHE 7
Db      362 AAADPHE 368
      |||||
Search completed: April 19, 2004, 11:51:20
Job time : 4.11727 secs
```

```
PT plasma expanders.
XX
XX Claim 2; Page 9; 20pp; English.
PS
CC N-terminal fragment of human serum albumin used as plasma expander, or as
CC substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
XX Sequence 388 AA;
SQ
    Query Match          100.0%; Score 38; DB 1; Length 388;
    Best Local Similarity 100.0%; Pred. No. 22;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAADPHE 7
Db      362 AAADPHE 368
      |||||
RESULT 15
AAP90391
ID AAP90391 standard; protein; 390 AA.
XX
XX AAP90391;
AC
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
XX N-terminal human serum albumin.
DE
XX
XX N-terminal portion of human serum albumin; plasma expanders.
KW
XX
XX Homo sapiens; (Human).
OS
XX
XX EP322094-A.
PN
XX 28-JUN-1989.
PD
XX 25-OCT-1988; 88EP-00310000.
PF
XX 30-OCT-1987; 87GB-00025529.
PR
XX (DELZ ) DELTA BIOTECHNOLOGY LTD.
PA
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
PI
XX WPI; 1989-186464/26.
DR
XX
XX New N-terminal fragments of human serum albumin - esp. useful as blood
PT plasma expanders.
FT
XX
XX Claim 2; Page 9; 20pp; English.
PS
XX N-terminal portion of human serum albumin. Used to make plasma expanders,
CC or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25
CC -MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
CC
XX
XX Sequence 389 AA;
SQ
    Query Match          100.0%; Score 38; DB 1; Length 389;
    Best Local Similarity 100.0%; Pred. No. 22;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAADPHE 7
Db      362 AAADPHE 368
      |||||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 2.94737 Seconds
(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_362_368

Perfect score: 38

Sequence: 1 AAADPHE 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	13	9	US-09-791-378-261
2	38	100.0	585	9	US-09-929-552-2
3	38	100.0	585	10	US-09-932-613-445
4	38	100.0	585	10	US-09-984-010-26
5	38	100.0	585	10	US-09-833-041-18
6	38	100.0	585	10	US-09-833-117-18
7	38	100.0	585	10	US-09-932-322-445
8	38	100.0	585	10	US-09-832-501-18
9	38	100.0	585	11	US-09-833-118-18
10	38	100.0	585	11	US-09-833-245-18
11	38	100.0	585	12	US-10-424-999-11
12	38	100.0	585	12	US-10-425-000-31
13	38	100.0	585	12	US-10-433-108-34
14	38	100.0	585	13	US-10-153-064-5
15	38	100.0	585	14	US-10-153-064A-5

16	38	100.0	585	14	US-10-319-263-1	Sequence 1, Appli
17	38	100.0	585	14	US-10-319-263-2	Sequence 2, Appli
18	38	100.0	585	14	US-10-414-469-1	Sequence 1, Appli
19	38	100.0	585	14	US-10-414-469-2	Sequence 2, Appli
20	38	100.0	585	14	US-10-413-831-1	Sequence 1, Appli
21	38	100.0	585	14	US-10-413-831-2	Sequence 2, Appli
22	38	100.0	585	15	US-10-413-832-1	Sequence 1, Appli
23	38	100.0	585	15	US-10-413-832-2	Sequence 2, Appli
24	38	100.0	585	15	US-10-414-386-1	Sequence 1, Appli
25	38	100.0	585	15	US-10-414-386-2	Sequence 2, Appli
26	38	100.0	585	15	US-10-233-675A-11	Sequence 11, Appli
27	38	100.0	585	15	US-10-462-262-26	Sequence 26, Appli
28	38	100.0	604	10	US-09-984-010-7	Sequence 7, Appli
29	38	100.0	609	10	US-09-919-039-370	Sequence 370, App
30	38	100.0	609	12	US-10-509-346-12	Sequence 12, Appli
31	38	100.0	609	13	US-10-153-064-7	Sequence 7, Appli
32	38	100.0	609	14	US-10-153-604A-7	Sequence 7, Appli
33	38	100.0	609	14	US-10-365-623-23	Sequence 23, Appli
34	38	100.0	610	9	US-09-984-186-2	Sequence 2, Appli
35	38	100.0	610	14	US-10-237-667-2	Sequence 2, Appli
36	38	100.0	610	14	US-10-237-708-2	Sequence 2, Appli
37	38	100.0	610	14	US-10-237-866-2	Sequence 2, Appli
38	38	100.0	610	14	US-10-237-871-2	Sequence 2, Appli
39	38	100.0	610	14	US-10-237-824-2	Sequence 13, Appli
40	38	100.0	616	12	US-10-433-108-13	Sequence 13, Appli
41	38	100.0	624	12	US-10-433-108-16	Sequence 16, Appli
42	38	100.0	631	12	US-10-433-108-14	Sequence 14, Appli
43	38	100.0	640	12	US-10-433-108-15	Sequence 15, Appli
44	38	100.0	640	12	US-10-433-108-17	Sequence 17, Appli
45	38	100.0	651	13	US-10-153-064-133	Sequence 133, App

ALIGNMENTS

RESULT 1

US-09-791-378-261
; Sequence 261, Application US/09791378
; Patent No. US20020142303A1
; GENERAL INFORMATION:
; APPLICANT: Parekh, Rajesh
; TITLE OF INVENTION: PROTEINS, GENES AND THEIR USE FOR DIAGNOSIS AND TREATMENT OF
; FILE REFERENCE: 9195-061-399
; CURRENT APPLICATION NUMBER: US/09/791,378
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/750,395
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 677
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 261
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-378-261

Query Match 100.0%; Score 38; DB 9; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.2; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

QY 1 AAADPHE 7
Db 3 AAADPHE 9

RESULT 2

US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Somnenschein, Carlos
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/929,552
FILING DATE: 14-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/769,746
FILING DATE: 19-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: MBRI-02584
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match 100.0%; Score 38; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
|||||
DB 362 AAADPHE 368

RESULT 3
US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT: DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-932-613-445

Query Match 100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
|||||

Db 362 AAADPHE 368

RESULT 4
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164
; FILING DATE: 19-DEC-1996
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
|||||
DB 362 AAADPHE 368

RESULT 5
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseitine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF545
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 18

```

; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match      100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
DB      362 AAADPHE 368

RESULT 6
US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match      100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
DB      362 AAADPHE 368

RESULT 7
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US2003019473A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT: DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-322-445

Query Match      100.0%; Score 38; DB 10; Length 585;
```

```

; Best Local Similarity 100.0%; Pred. No. 49;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
DB      362 AAADPHE 368

RESULT 8
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
DB      362 AAADPHE 368

RESULT 9
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Hasseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 38; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 AAADPHE 7
DB      362 AAADPHE 368

RESULT 10
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match      100.0%; Score 38; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
DB      362 AAADPHE 368

RESULT 11
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match      100.0%; Score 38; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
DB      362 AAADPHE 368

RESULT 12
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 38; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
DB      362 AAADPHE 368

RESULT 13
US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34

Query Match      100.0%; Score 38; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAADPHE 7
DB      362 AAADPHE 368

RESULT 14
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
```

; PRIOR APPLICATION NUMBER: 60/293,212
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-10-153-064-5

Query Match 100.0%; Score 38; DB 13; Length 585;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
 Db 362 AAADPHE 368

RESULT 15
 US-10-153-604A-5
 ; Sequence 5, Application US/10153604A
 ; Publication No. US20030143191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bell et al.
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PFS56
 ; CURRENT APPLICATION NUMBER: US/10/153,604A
 ; CURRENT FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: 60/293,212
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-10-153-604A-5

Query Match 100.0%; Score 38; DB 14; Length 585;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAADPHE 7
 Db 362 AAADPHE 368

Search completed: April 19, 2004, 12:54:59
 Job time : 2.94737 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 3.47368 Seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_280_288
Perfect score: 47
Sequence: 1 EKPLLEKSH 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL25:
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	Length	ID	Description
1	47	100.0	4	Q86YGO		Q86YGO homo sapien
2	39	83.0	584	Q7YSG3		Q7YSG3 felis silve
3	37	78.7	245	16 Q7QW3		Q7QW3 bordetella
4	37	78.7	245	16 Q7W1Y5		Q7W1Y5 bordetella
5	37	78.7	576	11 Q8C7C7		Q8C7C7 mus musculu
6	37	78.7	608	5 Q95VB7		Q95VB7 schistosoma
7	37	78.7	608	11 Q8C7H3		Q8C7H3 mus musculu
8	37	78.7	1381	10 Q94C01		Q94C01 hordeum vul
9	36	76.6	131	10 Q8LFB3		Q8LFB3 arabidopsis
10	36	76.6	480	16 Q7JQ53		Q7JQ53 rhodospirill
11	35	74.5	123	2 Q8KLA6		Q8KLA6 rhizobium e
12	35	74.5	156	16 Q97F20		Q97F20 clostridium
13	35	74.5	226	5 Q8I621		Q8I621 plasmodium
14	35	74.5	301	16 Q9PMW9		Q9PMW9 campylobact
15	35	74.5	1265	5 Q01495		Q01495 caenorhabdi
16	35	74.5	1380	10 Q40001		Q40001 hordeum vul

17	35	74.5	1534	5 Q8MPV7		Q8MPV7 caenorhabdi
18	35	74.5	1538	5 Q8MPV6		Q8MPV6 caenorhabdi
19	34	72.3	138	2 Q84B18		Q84B18 streptococc
20	34	72.3	191	16 Q81SB7		Q81SB7 bacillus an
21	34	72.3	191	16 Q81FA3		Q81FA3 bacillus ce
22	34	72.3	226	16 Q8ZS19		Q8ZS19 anabaena sp
23	34	72.3	244	2 Q34Z50		Q34Z50 wolinnella s
24	34	72.3	244	2 Q34Z50		Q34Z50 wolinnella s
25	34	72.3	250	4 Q9NS78		Q9NS78 homo sapien
26	34	72.3	265	4 Q9NS79		Q9NS79 homo sapien
27	34	72.3	269	4 Q9H4F5		Q9H4F5 homo sapien
28	34	72.3	269	11 Q9CQ04		Q9CQ04 mus musculu
29	34	72.3	269	11 Q9DAD7		Q9DAD7 mus musculu
30	34	72.3	311	2 Q86E51		Q86E51 haemophilus
31	34	72.3	477	16 Q87Y51		Q87Y51 pseudomonas
32	34	72.3	528	10 Q94FY3		Q94FY3 zea mays (m
33	34	72.3	560	13 Q7ZTT2		Q7ZTT2 brachydanio
34	34	72.3	586	3 Q9P985		Q9P985 saccharomyc
35	34	72.3	586	3 Q9P986		Q9P986 saccharomyc
36	34	72.3	586	3 Q9P984		Q9P984 saccharomyc
37	34	72.3	586	3 Q8J2S8		Q8J2S8 saccharomyc
38	34	72.3	747	5 Q7YSP6		Q7YSP6 caenorhabdi
39	34	72.3	768	2 Q9LBP8		Q9LBP8 ruminococcu
40	34	72.3	861	5 Q3XWF9		Q3XWF9 caenorhabdi
41	33	70.2	171	16 Q8ZQ00		Q8ZQ00 salmonella
42	33	70.2	173	16 Q98KG5		Q98KG5 rhizobium 1
43	33	70.2	195	4 Q9H855		Q9H855 homo sapien
44	33	70.2	270	16 Q7VK53		Q7VK53 helicobacte
45	33	70.2	271	16 Q8XLY6		Q8XLY6 clostridium

ALIGNMENTS

RESULT 1

Q86YGO PRELIMINARY; PRT; 417 AA.
AC Q86YGO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
EL Submitted (JSC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 2.
DR PRINTS; PR00302; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47360 MW; 16E764833EEF4B8D CRC64;

Query Match 100.0%; Score 47; DB 4; Length 417;

Best Local Similarity 100.0%; Pred. No. 0.5;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9

Db 112 EKPLLEKSH 120

RESULT 2

```

Q7YSG3
ID Q7YSG3 PRELIMINARY; PRT; 584 AA.
AC Q7YSG3;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Albumin (Fragment).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reiniger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
RT albumin: IGE recognition, induction of basophil activation and
RT lymphoproliferative responses in atopic patients.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487677; CAD22275.1; -.
FT NON_TER
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 83.0%; Score 39; DB 6; Length 584;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KPLLEKSH 9
Db 281 KPVLEKSH 288

RESULT 3
Q7WQW3 PRELIMINARY; PRT; 245 AA.
AC Q7WQW3;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Phosphorylase family protein.
GN BB0212.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;
"Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; EX640437; CAE30710.1; -.
KW Complete proteome.
SQ SEQUENCE 245 AA; 25969 MW; 66554503B72F87 CRC64;

Query Match 78.7%; Score 37; DB 16; Length 245;
Best Local Similarity 77.8%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9
Db 110 EYPLLEKSH 118

RESULT 5
Q8C7C7 PRELIMINARY; PRT; 576 AA.
ID Q8C7C7;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=1246851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.

```

```

Db 110 EYPLLEKSH 118
| |||||:|
| |||||:|

RESULT 4
Q7WLY5 PRELIMINARY; PRT; 245 AA.
ID Q7WLY5;
AC Q7WLY5;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Phosphorylase family protein.
GN BPP0208.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;
"Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; EX640423; CAE39949.1; -.
KW Complete proteome.
SQ SEQUENCE 245 AA; 25999 MW; 3B2924005602E7EE CRC64;

Query Match 78.7%; Score 37; DB 16; Length 245;
Best Local Similarity 77.8%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EKPLLEKSH 9
Db 110 EYPLLEKSH 118
| |||||:|
| |||||:|

RESULT 5
Q8C7C7 PRELIMINARY; PRT; 576 AA.
ID Q8C7C7;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=1246851;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.

```

```

DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR DR PDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER
SQ SEQUENCE 576 AA; 65002 MW; F85733B99AE37F04 CRC64;

Query Match 78.7%; Score 37; DB 11; Length 576;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
:|||||:|
Db 272 DKPLLEKSH 280

RESULT 6
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
SEQUENCE FROM N.A.
RA Oman A., Asahi H., Stadecker M.J., Loverde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
component in murine infection with Schistosoma mansoni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF18550; AL08579.1; -.
DR GO: GO:0005615; C:extracellular space; IEA.
DR GO: GO:0005386; F:carrier activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68225 MW; E5EAB28E1C66E54 CRC64;

Query Match 78.7%; Score 37; DB 5; Length 608;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
:|||||:|
Db 304 DKPLLEKSH 312

RESULT 7
Q8C7H3 PRELIMINARY; PRT; 608 AA.
AC Q8C7H3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin 1.
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.

```

```

RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT Nature 420:563-573(2002).";
RL Nature 420:563-573(2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI:87991; Alb1.
DR GO: GO:0005615; C:extracellular space; IEA.
DR GO: GO:0005386; F:carrier activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;

Query Match 78.7%; Score 37; DB 11; Length 608;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPLLEKSH 9
:|||||:|
Db 304 DKPLLEKSH 312

RESULT 8
Q94C01 PRELIMINARY; PRT; 1381 AA.
ID Q94C01
AC Q94C01;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mg-chelatase subunit XANTHA-F.
GN XANTHA-F.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Svalcof's Bonus;
RA Olsson U.;
RT "Barley (Hordeum vulgare) Mg-chelatase subunit (Xantha-f) gene.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY039003; AAK72401.1; -.
DR GO: GO:0009058; P:biosynthesis; IEA.
DR InterPro: IPR003672; CobN/Mg chltase.
DR Pfam; PF02514; cobN-Mg_chel; 1.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 1381 AA; 152900 MW; ED91EA6CFOF23B5B CRC64;

Query Match 78.7%; Score 37; DB 10; Length 1381;
Best Local Similarity 87.5%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KPLLEKSH 9
:|||||:|
Db 465 KPLLEKSH 472

RESULT 9
Q8LFB3 PRELIMINARY; PRT; 191 AA.
ID Q8LFB3
AC Q8LFB3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation."; 0:0-0(2002).
 RL Genome Biol. 0:0-0(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY084948; AAM61509.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 191 AA; 21908 MW; F7434F97C7520294 CRC64;
 Query Match 76.6%; Score 36; DB 10; Length 191;
 Best Local Similarity 77.8%; Pred. No. 37;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 EXPLEKSH 9
 Db 151 EXPLEKKN 159
 RESULT 10
 Q7UQ53 PRELIMINARY; PRT; 480 AA.
 AC Q7UQ53;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN RB6510.
 OS Rhodospirillum rubrum.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Fireillula.
 OX NCBI_TaxID=117;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416;
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Bortym K., Heitmann K., Rabus R.,
 RA Schlesner H., Amann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Fireillula sp.
 RT strain 1."; 0:0-0(2003).
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
 DR EMBL: BX294144; CAD74852.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 480 AA; 53002 MW; 82507EAS798FB7A2 CRC64;
 Query Match 76.6%; Score 36; DB 16; Length 480;
 Best Local Similarity 87.5%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 KPLEKSH 9
 Db 66 KPLEKHLH 73
 RESULT 11
 Q8KLA6 PRELIMINARY; PRT; 123 AA.
 AC Q8KLA6;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Hypothetical protein.
 GN YP017.
 OS Rhizobium etli.
 OC Plasmid symbiotic plasmid P42d.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
 OX NCBI_TaxID=29449;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CFN42;
 RX MEDLINE=91193195; PubMed=2013564;
 RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;
 RT "Structural complexity of the symbiotic plasmid of Rhizobium
 RT leguminosarum bv. phaseoli."; 0:0-0(1991).
 RL J. Bacteriol. 173:2411-2419(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=CFN42;
 RX MEDLINE=97419521; PubMed=9274036;
 RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,
 RA Cevallos M.A., Davila G.;
 RT "Sequence, localization and characteristics of the replicator region
 RT of the symbiotic plasmid of Rhizobium etli."; 0:0-0(1997).
 RL Microbiology 143:2825-2831(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CFN42;
 RA Quintero V., Cevallos M.A., Davila G.;
 RT "A site-specific recombinase and RecA are required to exert
 RT incompatibility towards the symbiotic plasmid of Rhizobium etli."; 0:0-0(1997).
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U80928; AAM54841.1; -;
 DR GO: 0046821; C:extrachromosomal DNA; IEA.
 KW Hypothetical protein; Plasmid.
 SQ SEQUENCE 123 AA; 13938 MW; 7349C06DC4D4766A CRC64;
 Query Match 74.5%; Score 35; DB 2; Length 123;
 Best Local Similarity 75.0%; Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 KPLEKSH 9
 Db 38 KPLEKRS 45
 RESULT 12
 Q97F20 PRELIMINARY; PRT; 156 AA.
 AC Q97F20;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Predicted acetyltransferase.
 GN CAC2935.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum"; 0:0-0(2001).
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL: AE007790; AAK80877.1; -;
 DR F1R; B97261; B97261.
 DR GO: 0008080; F:N-acetyltransferase activity; IEA.
 DR GO: 0016740; P:transferase activity; IEA.


```

DR InterPro; IPR000182; GCN5acetyl_trans.
KW Pfam; PF00583; Acetyltransf; 1.
KW Transferrase; Complete proteome.
SQ SEQUENCE 156 AA; 17998 MW; ADP750D0FC593E8 CRC64;

Query Match 74.5%; Score 35; DB 15; Length 156;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPILLEKSH 9
   |||||
Db 137 EKPILFETH 145

RESULT 13
ID Q81621 PRELIMINARY; PRT; 226 AA.
AC Q81621;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cyclophilin, putative.
GN PF0120C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum".
RL Nature 419:498-511(2002).
DR EMBL; AE014844; AAN36113.1; -.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR002130; CSA.PPIase.
DR Pfam; PF00160; pro.isomerase; 1.
DR PRINTS; PR00153; CSAPPISMRASE.
DR PROSITE; PS50072; CSA.PPIase_2; 1.
SQ SEQUENCE 226 AA; 26429 MW; 3A5E790BB457E181 CRC64;

Query Match 74.5%; Score 35; DB 5; Length 226;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EKPILLEK 7
   |||||
Db 197 EKPILLEK 203

RESULT 14
ID Q9PMW9 PRELIMINARY; PRT; 301 AA.
AC Q9PMW9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Cj1330.
GN Cj1330.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Waterston R.;
RT "Direct Submission".
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; U97407; AAB52482.2; -.
DR PIR; D87789; D87789.
DR WormPep; C34G6.4; CE29212.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD0000006; ABC_transporter; 2.

```

Mon Apr 19 13:27:19 2004

DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
KW Hypothetical protein; ATP-binding; Transport
SQ SEQUENCE 1265 AA; 140465 MW; 4948EF5C5A402757 CRC64;
Query Match 74.5%; Score 35; DB 5; Length 1265;
Best Local Similarity 75.0%; Pred. NO. 3.7e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 KPLLEKSH 9
Db 13 KPLLEKSH 20

Search completed: April 19, 2004, 12:00:09
Job time : 5.47368 secs

```

DR InterPro; IPR006140; 2-Hacid DH_C.
DR Pfam; PF00389; 2-Hacid DH; 1.
DR Pfam; PF02826; 2-Hacid DH_C; 1.
DR KMW plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 315 AA; 34163 MW; A3357B4D7608BE26 CRC64;

Query Match          90.6%; Score 29; DB 16; Length 315;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
DB 31 EAADKAA 37

RESULT 11
ID Q7U021 PRELIMINARY; PRT; 387 AA.
AC Q7U021;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Probable succinyl-CoA synthetase (Beta chain) SUCC (SCS-beta)
DE (EC 6.2.1.5).
DE SUCC OR MB0976.
GN Mycobacterium bovis.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Ackin R., Doggett J., Mayes R., Keating M., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RA "The complete genome sequence of Mycobacterium bovis.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248337; CAD93837.1; -.
DR Ligase; Complete proteome.
KW Ligase; Complete proteome.
SQ SEQUENCE 387 AA; 40897 MW; DC055A7CD5E987CC CRC64;

Query Match          90.6%; Score 29; DB 16; Length 387;
Best Local Similarity 85.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
DB 376 EAADKAA 382

RESULT 12
ID Q82PD6 PRELIMINARY; PRT; 466 AA.
AC Q82PD6;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN SAV987.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Omura S., Takahashi Y., Horioka H., Nakazawa H., Geonice T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
metabolites.";
Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
Proc. Natl. Acad. Sci. U.S.A. 100:526-531(2003).
DR EMBL; AP005025; BAC68697.1; -.
DR Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 466 AA; 48152 MW; 436EEF5AC2E3FD2 CRC64;

Query Match          90.6%; Score 29; DB 16; Length 466;
Best Local Similarity 85.7%; Pred. No. 4.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
DB 60 EAADKAA 66

RESULT 13
ID Q8NST2 PRELIMINARY; PRT; 503 AA.
AC Q8NST2;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein Cg10585.
GN CGL0585.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005275; BAB97978.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 503 AA; 52191 MW; 7C257A5457A515C1 CRC64;

Query Match          90.6%; Score 29; DB 16; Length 503;
Best Local Similarity 85.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
DB 106 EAADKAA 112

RESULT 14
ID Q8XN8 PRELIMINARY; PRT; 574 AA.
AC Q8XN8;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE LD23509P (CG12141-PA).
GN RATS-LYS OR CG12141.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Drensek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides F.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA April J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Bavendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foler C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gorg F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Haritz N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.L., Ibegwam J.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Pallazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weissstock G.M., Weissbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong X., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Drensek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jallali M., Kruse D., Li P., Mattai B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,

RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome."
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY089547; AAL90285.1; -;
DR EMBL; AE003447; AAN09255.1; -;
DR FlyBase; FB00027084; Aats-lys.
DR GO; GO:0005737; F: Cytoplasm; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0004824; F: lysine-tRNA ligase activity; IEA.
DR GO; GO:0003676; F: nucleic acid binding; IEA.
DR GO; GO:0006430; P: lysyl-tRNA aminoacylation; IEA.
DR InterPro; IPR008994; Nucleic acid OB.
DR InterPro; IPR004364; tRNA-synt_2_OB.
DR InterPro; IPR002313; tRNA-synt_lys_2.
DR InterPro; IPR004365; tRNA-anti.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF00152; tRNA-synt_2; 1.
DR Pfam; PF01336; tRNA-anti; 1.
DR PRINTS; PR00982; TRNASYNTHLYS.
DR TIGRFAMs; TIGR00499; lysS_bact; 1.
DR PROSITE; PS00862; AT TRNA LIGASE II; 1.
DR SEQUENCE 574 AA; 64661 MW; 861DAEA9C53E0DBE CRC64;
SQ
Query Match 90.6%; Score 29; DB 5; Length 574;
Best Local Similarity 85.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQKAA 7
DB 498 QASDKAA 504

RESULT 15
QYSG3 PRELIMINARY; PRT; 584 AA.
ID Q7YSG3
AC Q7YSG3
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin (Fragment).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=livar;
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
RT albumin: IGE recognition, induction of basophil activation and
RT lymphoproliferative responses in atopic patients."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487677; CAD32275.1; -;
FT NON_TER 1
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;
Query Match 90.6%; Score 29; DB 6; Length 584;
Best Local Similarity 85.7%; Pred. No. 6.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQKAA 7
DB 498 QASDKAA 504

Db 170 EAADKAA 176

Search completed: April 19, 2004, 12:00:05
Job time : 4.70175 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 4.11727 Seconds
(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_170_176

Perfect score: 32
Sequence: 1 QADKAA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	116	4	AAO00108 Human pol
2	32	100.0	143	4	AAO02571 Human pol
3	32	100.0	192	4	AAU29875 Novel hum
4	32	100.0	195	5	AAO17048 Human ser
5	32	100.0	195	7	ABU10022 Human ser
6	32	100.0	204	3	AAy83947 Yeast cod
7	32	100.0	214	4	AAU29874 Novel hum
8	32	100.0	236	5	AAO17051 Human alb
9	32	100.0	241	5	AAO16984 Alpha-MSH
10	32	100.0	241	7	ABU10025 Alpha-MSH
11	32	100.0	242	5	AAO16985 Alpha-MSH
12	32	100.0	244	5	AAO16986 Alpha-MSH
13	32	100.0	245	5	AAO16987 Alpha-MSH
14	32	100.0	245	5	AAO16988 Alpha-MSH
15	32	100.0	268	5	AAO16989 Alpha-MSH
16	32	100.0	268	7	ABU10026 Human ser
17	32	100.0	303	2	AAr14178 Human ser
18	32	100.0	327	4	AAU32564 Novel hum
19	32	100.0	327	4	AAU29942 Novel hum
20	32	100.0	327	4	AAU32995 Novel hum
21	32	100.0	327	4	AAU33287 Novel hum
22	32	100.0	373	1	AAp90387 N-termina
23	32	100.0	388	1	AAp90389 N-termina
24	32	100.0	389	1	AAp90390 N-termina
25	32	100.0	390	1	AAp90391 N-termina

26	32	100.0	401	4	AAU29876 Novel hum
27	32	100.0	407	1	AAp90392 N-termina
28	32	100.0	463	2	AAr14179 Human ser
29	32	100.0	500	7	ADD32019 Heterolog
30	32	100.0	550	4	AAU29877 Novel hum
31	32	100.0	584	6	ABG72381 Mature hu
32	32	100.0	585	1	AAp93344 Sequence
33	32	100.0	585	1	AAp90388 Mature hu
34	32	100.0	585	1	AAp91422 Human nor
35	32	100.0	585	1	AAr05318 Human ser
36	32	100.0	585	2	AAr08457 Human ser
37	32	100.0	585	2	AAr26207 Human ser
38	32	100.0	585	2	AAr26362 Synthetic
39	32	100.0	585	2	AAr20029 Human ser
40	32	100.0	585	2	AAr80301 Human ser
41	32	100.0	585	2	AAO20111 HSA prote
42	32	100.0	585	2	AAw59841 Mature pr
43	32	100.0	585	3	AAy84873 Amino aci
44	32	100.0	585	3	AAy83946 Yeast cod
45	32	100.0	585	4	AAw52567 Mature hu

ALIGNMENTS

RESULT 1
AAO00108
ID AAO00108 standard; protein; 116 AA.
XX AC AAO00108;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 14000.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.
XX PN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX PS WPI; 2001-514838/56.
XX PS N-PSDB; AAI50039.
XX CC Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX CC and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 20; SEQ ID NO 14000; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 116 AA;

Query Match 100.0%; Score 32; DB 4; Length 116;
 Best Local Similarity 100.0%; Pred. No. 24; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0;

QY 1 QAADKAA 7
 DB 11 QAADKAA 17
 |||||

RESULT 2
 AA002571
 ID AA002571 standard; protein; 143 AA.
 XX
 AC AA002571;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 16463.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004927.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 PI WPI; 2001-514838/56.
 DR N-PSDB; AAI82502.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 XX
 PS Claim 20; SEQ ID NO 16463; 1399pp + Sequence Listing; English.
 CC
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 143 AA;

Query Match 100.0%; Score 32; DB 4; Length 143;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
 DB 11 QAADKAA 17
 |||||

RESULT 3
 AAU29875
 ID AAU29875 standard; protein; 192 AA.
 XX
 AC AAU29875;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #366.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US008656.
 XX
 PR 18-APR-2000; 2000US-00552929.
 PR 26-JAN-2001; 2001US-00770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 PI WPI; 2001-611725/70.
 DR
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX
 PS Claim 20; Page 206; 765pp; English.
 CC
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 192 AA;

Query Match 100.0%; Score 32; DB 4; Length 192;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
 DB 176 QAADKAA 182
 |||||

RESULT 4
 AA017048
 ID AA017048 standard; protein; 195 AA.


```

overlapping oligonucleotide; expression vector.
XX Homo sapiens.
OS Synthetic.
XX
XX CN1239103-A.
XX
XX 22-DEC-1999.
XX
XX 17-JUN-1998; 98CN-00102506.
XX
XX 17-JUN-1998; 98CN-00102506.
XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.
XX
XX Li S, Lu D;
PI
XX WPI; 2000-351198/31.
DR N-PSDB; AAA10092.
XX
XX Process for preparing recombinant human serum albumin comprising yeast
PT biased sex codons - uses a recombinant DNA technique.
XX
XX Example 1; Fig 3; 44pp; Chinese.
XX
XX The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
CC three synthetic fragments (AAA10092-A10094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the sequence of the HSA fragment HSA-I encoded by the human gene with a
CC yeast codon bias. The invention also covers a recombinant expression
CC vector, yeast host cells carrying the recombinant expression vector and
CC the process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
XX Sequence 204 AA;
SQ
Query Match 100.0%; Score 32; DB 3; Length 204;
Best Local Similarity 100.0%; Pred No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 QAADKAA 7
Db 177 QAADKAA 183
|||||
RESULT 7
AAU29874
ID AAU29874 standard; protein; 214 AA.
AC
XX AAU29874;
XX
XX 18-DEC-2001 (first entry)
DT
XX
XX Novel human secreted protein #365.
DE
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
OS
XX WO200179449-A2.
XX
XX 25-OCT-2001.
PD
XX
XX 16-APR-2001; 2001WO-US008656.
XX
XX 18-APR-2000; 2000US-00552929.
XX
XX 26-JAN-2001; 2001US-00770160.
XX

```

XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX WPI; 2002-195801/25.
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX Example 2; Page 48; 89pp; English.
XX The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a peptide described in the exemplification of the
CC invention
XX Sequence 236 AA;
SQ Query Match 100.0%; Score 32; DB 5; Length 236;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAADKAA 7
DB 194 QAADKAA 200
RESULT 9
AAO16984
ID AAO16984 standard; protein; 241 AA.
XX AAO16984;
AC AAO16984;
DT 29-MAY-2002 (first entry)
XX Alpha-MSH construct protein fragment SEQ ID NO: 59.
DE Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
XX Alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; anti-inflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX Unidentified.
OS WO200206316-A2.
XX 24-JAN-2002.
PN 16-JUL-2001; 2001WO-US022263.
PD 14-JUL-2000; 2000US-0218381P.
XX 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-023380P.
XX 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX (ZYCO-) ZYCOs INC.
PA Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX WPI; 2002-195801/25.
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory

PT or autoimmune disorders.
XX Example 2; Page 4-5; 89pp; English.
XX The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX Sequence 241 AA;
SQ Query Match 100.0%; Score 32; DB 5; Length 241;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAADKAA 7
DB 194 QAADKAA 200
RESULT 10
ABU10025
ID ABU10025 standard; protein; 241 AA.
XX ABU10025;
AC ABU10025;
XX 31-JUL-2003 (first entry)
DT Alpha-MSH/Serum albumin fusion protein H9.
DE Bladder disorder; cytostatic; anti-inflammatory; immune response;
XX un-methylated CpG sequence; alpha-MSH; melanocortin receptor;
KW bladder cancer; tumour; interstitial cystitis; inflammation;
KW alpha-MSH concatamer; melanocyte stimulating hormone; human;
KW serum albumin; fusion protein; H9.
XX Homo sapiens.
OS Synthetic.
OS US2002193332-A1.
PN 19-DEC-2002.
XX 12-FEB-2002; 2002US-00074956.
PF 12-FEB-2001; 2001US-0268175P.
XX (HEDL/) HEDLEY M L.
PR Hedley ML;
XX WPI; 2003-447327/42.
DR Modulating immune responses in a mammal with a bladder disorder e.g.
XX bladder cancer, by administering nucleic acids comprising un-methylated
XX CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to
XX the mammal.
XX Example 2; Page 9; 17pp; English.
XX The invention describes a method of modulating an immune response in a
CC mammal, comprising identifying a mammal that has or is at risk for having
CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)
CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated
CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;
CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The
CC method is useful for modulating immune response in a mammal having a

CC bladder disorder, where administration of (N1) results in an amelioration
CC of one or more symptoms of the disorder. Preferably, the method is useful
CC for modulating immune response in a mammal having bladder cancer (where
CC administration of (N1) results in a decrease in tumour size or activity),
CC or for modulating immune response in a mammal having interstitial
CC cystitis (where administration of (N1) results in a modulation of the
CC immune response from Th2 response to a Th1 response). The method is also
CC useful for modulating immune response in a mammal having bladder disorder
CC that is characterised by inflammation which is associated with symptoms
CC of interstitial cystitis or associated with a disruption of the integrity
CC of the bladder lining. This is the amino acid sequence of a human serum
CC melanocyte stimulating hormone (alpha-MSH) concatamer useful in the
CC treatment of bladder disorders
XX
SQ Sequence 241 AA;

Query Match 100.0%; Score 32; DB 7; Length 241;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
| | | | |
DB 194 QAADKAA 200

RESULT 11
AAO16985
ID AAO16985 standard; protein; 242 AA.

XX AC AAO16985;
XX DT 29-MAY-2002 (first entry)
XX DE Alpha-MSH construct protein fragment SEQ ID NO: 70.
XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antinflammatory; antirheumatic; antiarthritic;
KW antidiabetic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.

XX OS Unidentified.

XX PN WO200206316-A2.

XX PD 24-JAN-2002.

XX PF 16-JUL-2001; 2001WO-US022263.

XX PR 14-JUL-2000; 2000US-0218381P.

XX PR 18-AUG-2000; 2000US-0228382P.

XX PR 06-OCT-2000; 2000US-0238380P.

XX PR 29-DEC-2000; 2000US-0258764P.

XX PR 14-JUN-2001; 2001US-0298317P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;

XX DR WPI; 2002-195801/25.

XX PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
XX stimulating hormone concatamer or its analog, for treating inflammatory
XX or autoimmune disorders.

XX PS Disclosure; Page 5; 89pp; English.

XX CC The present invention relates to a nucleic acid comprising a sequence
XX encoding a fusion polypeptide having an alpha-melanocyte stimulating
XX hormone (MSH) concatamer. The sequences are useful for treating an
XX individual suffering from, or at risk of, a disorder of the immune system

CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX

SQ Sequence 242 AA;

Query Match 100.0%; Score 32; DB 5; Length 242;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
| | | | |
DB 194 QAADKAA 200

RESULT 12

AAO16986
ID AAO16986 standard; protein; 244 AA.

XX AC AAO16986;

XX DT 29-MAY-2002 (first entry)

XX DE Alpha-MSH construct protein fragment SEQ ID NO: 71.

XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antinflammatory; antirheumatic; antiarthritic;
KW antidiabetic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.

XX OS Unidentified.

XX PN WO200206316-A2.

XX PD 24-JAN-2002.

XX PF 16-JUL-2001; 2001WO-US022263.

XX PR 14-JUL-2000; 2000US-0218381P.

XX PR 18-AUG-2000; 2000US-0228382P.

XX PR 06-OCT-2000; 2000US-0238380P.

XX PR 29-DEC-2000; 2000US-0258764P.

XX PR 14-JUN-2001; 2001US-0298317P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;

XX DR WPI; 2002-195801/25.

XX PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
XX stimulating hormone concatamer or its analog, for treating inflammatory
XX or autoimmune disorders.

XX PS Disclosure; Page 5; 89pp; English.

XX CC The present invention relates to a nucleic acid comprising a sequence
XX encoding a fusion polypeptide having an alpha-melanocyte stimulating
XX hormone (MSH) concatamer. The sequences are useful for treating an
XX individual suffering from, or at risk of, a disorder of the immune system
XX e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
XX arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
XX hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
XX multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
XX present sequence is a protein described in the exemplification of the
XX invention

SQ Sequence 244 AA;
 Query Match 100.0%; Score 32; DB 5; Length 244;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
 |||||
 Db 194 QAADKAA 200

RESULT 13
 AAO16987
 ID AAO16987 standard; protein; 245 AA.
 XX AC AAO16987;
 XX DT 29-MAY-2002 (first entry)
 XX DE Alpha-MSH construct protein fragment SEQ ID NO: 72.
 XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
 KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
 KW antidiabetic; antibacterial; dermatological; antipsoriatic;
 KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
 KW diabetes; uveitis; coeliac disease.
 XX OS Unidentified.
 XX PN WO200206316-A2.
 XX PD 24-JAN-2002.
 XX PF 16-JUL-2001; 2001WO-US022263.
 XX PR 14-JUL-2000; 2000US-0218381P.
 XX PR 18-AUG-2000; 2000US-0226382P.
 XX PR 06-OCT-2000; 2000US-0238380P.
 XX PR 29-DEC-2000; 2000US-0258764P.
 XX PR 14-JUN-2001; 2001US-0298317P.
 XX PA (ZYCO-) ZYCOS INC.
 XX PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
 XX WPI; 2002-195801/25.
 XX DR Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
 PT stimulating hormone concatamer or its analog, for treating inflammatory
 PT or autoimmune disorders.
 XX PS Disclosure; Page 5; 89pp; English.
 XX CC The present invention relates to a nucleic acid comprising a sequence
 CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
 CC hormone (MSH) concatamer. The sequences are useful for treating an
 CC individual suffering from, or at risk of, a disorder of the immune system
 CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
 CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
 CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
 CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
 CC present sequence is a protein described in the exemplification of the
 CC invention
 XX SQ Sequence 245 AA;
 Query Match 100.0%; Score 32; DB 5; Length 245;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
 |||||
 Db 194 QAADKAA 200

RESULT 15
 AAO16989
 ID AAO16989 standard; protein; 268 AA.

Db 194 QAADKAA 200
 |||||
 RESULT 14
 AAO16988
 ID AAO16988 standard; protein; 245 AA.
 XX AC AAO16988;
 XX DT 29-MAY-2002 (first entry)
 XX DE Alpha-MSH construct protein fragment SEQ ID NO: 73.
 XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
 KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
 KW antidiabetic; antibacterial; dermatological; antipsoriatic;
 KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
 KW diabetes; uveitis; coeliac disease.
 XX OS Unidentified.
 XX PN WO200206316-A2.
 XX PD 24-JAN-2002.
 XX PF 16-JUL-2001; 2001WO-US022263.
 XX PR 14-JUL-2000; 2000US-0218381P.
 XX PR 18-AUG-2000; 2000US-0226382P.
 XX PR 06-OCT-2000; 2000US-0238380P.
 XX PR 29-DEC-2000; 2000US-0258764P.
 XX PR 14-JUN-2001; 2001US-0298317P.
 XX PA (ZYCO-) ZYCOS INC.
 XX PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
 XX WPI; 2002-195801/25.
 XX DR Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
 PT stimulating hormone concatamer or its analog, for treating inflammatory
 PT or autoimmune disorders.
 XX PS Disclosure; Page 5; 89pp; English.
 XX CC The present invention relates to a nucleic acid comprising a sequence
 CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
 CC hormone (MSH) concatamer. The sequences are useful for treating an
 CC individual suffering from, or at risk of, a disorder of the immune system
 CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
 CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
 CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
 CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
 CC present sequence is a protein described in the exemplification of the
 CC invention
 XX SQ Sequence 245 AA;
 Query Match 100.0%; Score 32; DB 5; Length 245;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
 |||||
 Db 194 QAADKAA 200

RESULT 15
 AAO16989
 ID AAO16989 standard; protein; 268 AA.

Mon Apr 19 13:27:11 2004

XX AAO16989;
AC
XX 29-MAY-2002 (first entry)
DT
XX Alpha-MSH construct protein fragment SEQ ID NO: 60.
DE
XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antineoplastic; antirheumatic; antiarthritic;
KW antidiabetic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
OS Unidentified.
XX
XX WO200206316-A2.
FN
XX 24-JAN-2002.
PD
XX 16-JUL-2001; 2001WO-US022263.
PF
XX 14-JUL-2000; 2000US-0218381P.
PR 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-0238380P.
PR 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX
XX (ZYCO-) ZYCOS INC.
PA
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
FI WPI; 2002-195801/25.
XX
DR Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
XX Example 2; Page 5; 89pp; English.
PS
XX The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX
XX Sequence 268 AA;
SQ
Query Match 100.0%; Score 32; DB 5; Length 268;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 QAADKAA 7
DB 194 QAADKAA 200
Search completed: April 19, 2004, 11:51:18
Job time : 5.11727 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 2.94737 Seconds

(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_170_176

Perfect score: 32

Sequence: 1 QAAADKAA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pdb:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pdb:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pdb:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pdb:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pdb:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pdb:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pdb:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_NEW_PUB.pdb:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pdb:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pdb:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pdb:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pdb:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pdb:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pdb:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	104	12	US-10-424-599-268254
2	32	100.0	195	13	US-10-074-956-24
3	32	100.0	241	13	US-10-074-956-27
4	32	100.0	268	13	US-10-074-956-28
5	32	100.0	585	9	US-09-929-552-2
6	32	100.0	585	10	US-09-932-613-445
7	32	100.0	585	10	US-09-984-010-26
8	32	100.0	585	10	US-09-833-041-18
9	32	100.0	585	10	US-09-833-117-18
10	32	100.0	585	10	US-09-932-322-445
11	32	100.0	585	10	US-09-833-501-18
12	32	100.0	585	11	US-09-833-118-18
13	32	100.0	585	11	US-09-833-245-18
14	32	100.0	585	12	US-10-424-999-11
15	32	100.0	585	12	US-10-425-000-31

16	32	100.0	585	12	US-10-433-108-34	Sequence 34, Appl
17	32	100.0	585	13	US-10-153-064-5	Sequence 5, Appl
18	32	100.0	585	14	US-10-153-604A-5	Sequence 5, Appl
19	32	100.0	585	14	US-10-319-263-1	Sequence 1, Appl
20	32	100.0	585	14	US-10-319-263-2	Sequence 2, Appl
21	32	100.0	585	14	US-10-414-469-1	Sequence 1, Appl
22	32	100.0	585	14	US-10-414-469-2	Sequence 2, Appl
23	32	100.0	585	14	US-10-413-831-1	Sequence 1, Appl
24	32	100.0	585	14	US-10-413-831-2	Sequence 1, Appl
25	32	100.0	585	15	US-10-413-832-1	Sequence 1, Appl
26	32	100.0	585	15	US-10-414-386-1	Sequence 1, Appl
27	32	100.0	585	15	US-10-414-386-2	Sequence 2, Appl
28	32	100.0	585	15	US-10-233-675A-11	Sequence 11, Appl
29	32	100.0	585	15	US-10-462-262-26	Sequence 26, Appl
30	32	100.0	604	10	US-09-984-010-7	Sequence 7, Appl
31	32	100.0	609	10	US-09-919-039-370	Sequence 370, Appl
32	32	100.0	609	12	US-10-609-346-12	Sequence 12, Appl
33	32	100.0	609	13	US-10-153-064-7	Sequence 7, Appl
34	32	100.0	609	14	US-10-153-604A-7	Sequence 7, Appl
35	32	100.0	609	14	US-10-365-623-23	Sequence 23, Appl
36	32	100.0	610	9	US-09-984-186-2	Sequence 2, Appl
37	32	100.0	610	14	US-10-237-667-2	Sequence 2, Appl
38	32	100.0	610	14	US-10-237-708-2	Sequence 2, Appl
39	32	100.0	610	14	US-10-237-866-2	Sequence 2, Appl
40	32	100.0	610	14	US-10-237-871-2	Sequence 2, Appl
41	32	100.0	610	14	US-10-237-624-2	Sequence 13, Appl
42	32	100.0	616	12	US-10-433-108-13	Sequence 16, Appl
43	32	100.0	624	12	US-10-433-108-16	Sequence 16, Appl
44	32	100.0	631	12	US-10-433-108-14	Sequence 14, Appl
45	32	100.0	631	12	US-10-433-108-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-10-424-599-268254
; Sequence 268254, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 268254
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_84257C.1.pap
US-10-424-599-268254

Query Match 100.0%; Score 32; DB 12; Length 104;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QAAADKAA 7
Db 47 QAAADKAA 53

RESULT 2
US-10-074-956-24
; Sequence 24, Application US/10074956
; Publication No. US2002019332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne

;; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
;; FILE REFERENCE: 08191-022001
;; CURRENT APPLICATION NUMBER: US/10/074,956
;; PRIOR FILING DATE: 2002-06-10
;; PRIOR APPLICATION NUMBER: 60/268,175
;; PRIOR FILING DATE: 2001-02-12
;; NUMBER OF SEQ ID NOS: 29
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 24
;; LENGTH: 195
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-074-956-24

Query Match 100.0%; Score 32; DB 13; Length 195;
Best Local Similarity 100.0%; Pred. NO. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 3
US-10-074-956-27
;; Sequence 27, Application US/10074956
;; Publication No. US2002019332A1
;; GENERAL INFORMATION:
;; APPLICANT: Hedley, Mary Lynne
;; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
;; FILE REFERENCE: 08191-022001
;; CURRENT APPLICATION NUMBER: US/10/074,956
;; CURRENT FILING DATE: 2002-06-10
;; PRIOR APPLICATION NUMBER: 60/268,175
;; PRIOR FILING DATE: 2001-02-12
;; NUMBER OF SEQ ID NOS: 29
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 27
;; LENGTH: 241
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-074-956-27

Query Match 100.0%; Score 32; DB 13; Length 241;
Best Local Similarity 100.0%; Pred. NO. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7
Db 194 QAADKAA 200

RESULT 4
US-10-074-956-28
;; Sequence 28, Application US/10074956
;; Publication No. US2002019332A1
;; GENERAL INFORMATION:
;; APPLICANT: Hedley, Mary Lynne
;; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
;; FILE REFERENCE: 08191-022001
;; CURRENT APPLICATION NUMBER: US/10/074,956
;; CURRENT FILING DATE: 2002-06-10
;; PRIOR APPLICATION NUMBER: 60/268,175
;; PRIOR FILING DATE: 2001-02-12
;; NUMBER OF SEQ ID NOS: 29
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 28
;; LENGTH: 268
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-074-956-28

Query Match 100.0%; Score 32; DB 13; Length 268;

Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7
Db 194 QAADKAA 200

RESULT 5
US-09-929-552-2
;; Sequence 2, Application US/09929552
;; Patent No. US20020123080A1
;; GENERAL INFORMATION:
;; APPLICANT: Sonnenschein, Carlos
;; Soto, Ana M.
;; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Medlen & Carroll, LLP
;; STREET: 220 Montgomery Street, Suite 2200
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94104
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/929,552
;; FILING DATE: 14-Aug-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/769,746
;; FILING DATE: 19-DEC-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Carroll, Peter G.
;; REGISTRATION NUMBER: 32,837
;; REFERENCE/DOCKET NUMBER: MBRI-02584
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 585 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match 100.0%; Score 32; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. NO. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 6
US-09-932-613-445
;; Sequence 445, Application US/09932613
;; Publication No. US20030091565A1
;; GENERAL INFORMATION:
;; APPLICANT: Human Genome Sciences, Inc.
;; APPLICANT: Beltzer, James P.
;; APPLICANT: Potter, M. Daniel
;; APPLICANT: Fleming, Tony J.
;; APPLICANT: Rosen, Craig A.
;; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
;; FILE REFERENCE: Dyx-025.1 PCT; Dyx-025.1 US

```
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445

Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QAADKAA 7
Db      170 QAADKAA 176

RESULT 7
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164
; FILING DATE: 19-DEC-1996
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 QAADKAA 7
Db      170 QAADKAA 176

RESULT 8
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS45
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18
```

```
Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 QAADKAA 7
Db      170 QAADKAA 176
```

```
RESULT 9
US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS43
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18
```

```
Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 QAADKAA 7
Db      170 QAADKAA 176
```

```
RESULT 10
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
```



```

; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932.322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
; US-09-932-322-445

Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
      |||||
Db      170 QAADKAA 176

RESULT 11
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-832-501-18

Query Match      100.0%; Score 32; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
      |||||
Db      170 QAADKAA 176

RESULT 12
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; US-09-832-929-18_copy_170_176.rapb
```

```

; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-833-118-18

Query Match      100.0%; Score 32; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
      |||||
Db      170 QAADKAA 176

RESULT 13
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-833-245-18

Query Match      100.0%; Score 32; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
      |||||
Db      170 QAADKAA 176

RESULT 14
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for:
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
```

```
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match      100.0%; Score 32; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 15
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Camecon, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 32; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

Search completed: April 19, 2004, 12:54:58
Job time : 2.94737 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 2.14589 Seconds
(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_76_89

Perfect score: 70

Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	100.0	585	1	US-08-153-799-14
2	70	100.0	585	1	US-08-448-196A-3
3	70	100.0	585	2	US-08-984-176-1
4	70	100.0	585	2	US-08-702-572-2
5	70	100.0	585	3	US-08-769-746-2
6	70	100.0	585	4	US-10-153-064-5
7	70	100.0	609	1	US-08-223-619-3
8	70	100.0	609	1	US-08-433-037-4
9	70	100.0	609	4	US-08-897-956A-2
10	70	100.0	609	4	US-10-153-064-7
11	70	100.0	609	4	US-09-976-594-977
12	70	100.0	609	5	PCT-US95-04075-3
13	70	100.0	610	2	US-08-797-689-2
14	70	100.0	610	4	US-09-984-186-2
15	70	100.0	651	4	US-10-153-064-133
16	70	100.0	652	4	US-10-153-064-96
17	70	100.0	652	4	US-10-153-064-99
18	70	100.0	652	4	US-10-153-064-105
19	70	100.0	652	4	US-10-153-064-132
20	70	100.0	656	4	US-10-153-064-131
21	70	100.0	656	4	US-10-153-064-130
22	70	100.0	660	4	US-10-153-064-90
23	70	100.0	660	4	US-10-153-064-93
24	70	100.0	668	4	US-10-153-064-102
25	70	100.0	676	4	US-10-153-064-95
26	70	100.0	676	4	US-10-153-064-98
27	70	100.0	676	4	US-10-153-064-104

28	70	100.0	676	4	US-10-153-064-127	Sequence 127, App
29	70	100.0	676	4	US-10-153-064-129	Sequence 129, App
30	70	100.0	677	4	US-10-153-064-125	Sequence 125, App
31	70	100.0	680	4	US-10-153-064-123	Sequence 123, App
32	70	100.0	684	4	US-10-153-064-92	Sequence 92, Appl
33	70	100.0	692	4	US-10-153-064-101	Sequence 101, App
34	70	100.0	783	1	US-08-256-938-2	Sequence 2, Appli
35	70	100.0	787	1	US-08-256-938-4	Sequence 4, Appli
36	70	100.0	787	2	US-08-797-689-16	Sequence 16, Appl
37	70	100.0	787	4	US-09-984-186-1	Sequence 16, Appl
38	70	100.0	978	4	US-08-897-956A-3	Sequence 3, Appli
39	70	100.0	1184	4	US-10-153-064-89	Sequence 89, Appl
40	62	88.6	583	1	US-08-448-196A-6	Sequence 6, Appli
41	61	87.1	583	1	US-08-448-196A-5	Sequence 5, Appli
42	58	82.9	582	1	US-08-134-638-1	Sequence 1, Appli
43	58	82.9	583	1	US-08-448-196A-4	Sequence 4, Appli
44	42	60.0	584	1	US-08-448-196A-7	Sequence 7, Appli
45	40	57.1	104	4	US-09-134-000C-5225	Sequence 5225, Ap

ALIGNMENTS

RESULT 1
US-08-153-799-14
Sequence 14, Application US/08153799
Patent No. 5766883
GENERAL INFORMATION:
APPLICANT: Ballance, David J
APPLICANT: Goodey, Andrew R
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: R Hain Swope, BOC Health Care Inc
STREET: 100 Mountain Avenue
CITY: Murray Hill
STATE: New Jersey
COUNTRY: USA
ZIP: 07974
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Region
LOCATION: 369..419
OTHER INFORMATION: /note= "Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..585
OTHER INFORMATION: /note= "Amino acid sequence of
OTHER INFORMATION: natural HSA"

US-08-153-799-14

Query Match 100.0%; Score 70; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

RESULT 2

US-08-448-196A-3
Sequence 3, Application US/08448196A
Patent No. 5780594
GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C.
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
TITLE OF INVENTION: RELATED PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NASA
CITY: MARSHALL SPACE FLIGHT CENTER
STATE: ALABAMA
COUNTRY: USA
ZIP: 35812

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-448-196A-3

Query Match 100.0%; Score 70; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

RESULT 3

US-08-984-176-1
Sequence 1, Application US/08984176
Patent No. 5948609
GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C
APPLICANT: HO, JOSEPH X
APPLICANT: RUKER, FLORIAN
TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
FILE REFERENCE: 08/984,176
CURRENT APPLICATION NUMBER: US/08/984,176
CURRENT FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 1
LENGTH: 585
TYPE: PRT
ORGANISM: Homo sapiens
US-08-984-176-1

Query Match 100.0%; Score 70; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

RESULT 4

US-08-702-572-2
Sequence 2, Application US/08702572
Patent No. 5965386
GENERAL INFORMATION:

APPLICANT: Kerry-Williams, Sean M
APPLICANT: Gilbert, Sarah C
TITLE OF INVENTION: Yeast Strains and Modified Albumins
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Centeon L.L.C.
STREET: 1020 First Avenue
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-1310

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,572
FILING DATE: 11-NOV-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/23857
FILING DATE: 1-MAR-1995
APPLICATION NUMBER: GB 9404270.2
FILING DATE: 5-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Naomi Biswas
REGISTRATION NUMBER: 38,384
REFERENCE/DOCKET NUMBER: CE0114 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610/878/4294
TELEFAX: 610/878/4221
INFORMATION FOR SEQ ID NO: 2:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2
Query Match      100.0%; Score 70; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      76 TVATLRETYGEMAD 89

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MERRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2
Query Match      100.0%; Score 70; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      76 TVATLRETYGEMAD 89

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5
Query Match      100.0%; Score 70; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      76 TVATLRETYGEMAD 89

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3
Query Match      100.0%; Score 70; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      100 TVATLRETYGEMAD 113

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
```

APPLICANT: Barr, Kathryn A.
APPLICANT: Briarley, Russell A.
APPLICANT: Thill, Gregory P.
APPLICANT: Tschopp, Juerg F.
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
TITLE OF INVENTION: PICHIA PASTORIS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,037
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 91082
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 100.0%; Score 70; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 100 TVATLRETYGEMAD 113

RESULT 9
US-08-897-956A-2
Sequence 2, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/CPA
CURRENT APPLICATION NUMBER: US/08/897,956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 609
TYPE: PRT
ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 100.0%; Score 70; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 100 TVATLRETYGEMAD 113

RESULT 10
US-10-153-064-7
Sequence 7, Application US/10153064
Patent No. 6663485
GENERAL INFORMATION:
APPLICANT: Bell et al.
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
FILE REFERENCE: PFS56
CURRENT APPLICATION NUMBER: US/10/153,064
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: 60/293,212
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 137
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 609
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match 100.0%; Score 70; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 100 TVATLRETYGEMAD 113

RESULT 11
US-09-976-594-977
Sequence 977, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 977
LENGTH: 609
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match 100.0%; Score 70; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 100 TVATLRETYGEMAD 113

RESULT 12
PCT-US95-04075-3
Sequence 3, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.

;; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
;; TITLE OF INVENTION: Protein
;; NUMBER OF SEQUENCES: 33
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Amgen Center, Patent Operations/RRC
;; STREET: 1840 DeHavilland Drive
;; CITY: Thousand Oaks
;; STATE: California
;; COUNTRY: U.S.
;; ZIP: 91320-1789
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/04075
;; FILING DATE:
;; CLASSIFICATION:
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 609 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
PCT-US95-04075-3

Query Match 100.0%; Score 70; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 100 TVATLRETYGEMAD 113

RESULT 13
US-08-797-689-2
;; Sequence 2, Application US/08797689
;; Patent No. 5976969
;; GENERAL INFORMATION:
;; APPLICANT: Fleer, Reinhard
;; APPLICANT: Fournier, Alain
;; APPLICANT: Guittion, Jean-Dominique
;; APPLICANT: Jung, Gerard
;; APPLICANT: Yeh, Patrice
;; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;; CONTAINING SAID POLYPEPTIDES
;; NUMBER OF SEQUENCES: 36
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Rhone-Poulenc Rorer Inc.
;; STREET: 500 Arcola Road, 3C43
;; CITY: Collegeville
;; STATE: PA
;; COUNTRY: USA
;; ZIP: 19426
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Macintosh
;; OPERATING SYSTEM: System 7.1
;; SOFTWARE: Word 5.1 (Patentin)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/797,689
;; FILING DATE: 31-JAN-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/256,927
;; FILING DATE: 28-JUL-1994
;; APPLICATION NUMBER: FR 92/01064
;; FILING DATE: 31-JAN-1992
;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/FR93/00085
;; FILING DATE: 28-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith Ph.D., Julie K.
;; REGISTRATION NUMBER: P-38,619
;; REFERENCE/DOCKET NUMBER: ST92006-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (610) 454-3839
;; TELEFAX: (610) 454-3808
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 610 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-797-689-2

Query Match 100.0%; Score 70; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 100 TVATLRETYGEMAD 113

RESULT 14
US-09-984-186-2
;; Sequence 2, Application US/09984186
;; Patent No. 6686179
;; GENERAL INFORMATION:
;; APPLICANT: Fleer, Reinhard
;; APPLICANT: Fournier, Alain
;; APPLICANT: Guittion, Jean-Dominique
;; APPLICANT: Jung, Gerard
;; APPLICANT: Yeh, Patrice
;; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;; CONTAINING SAID POLYPEPTIDES
;; NUMBER OF SEQUENCES: 36
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Rhone-Poulenc Rorer Inc.
;; STREET: 500 Arcola Road, 3C43
;; CITY: Collegeville
;; STATE: PA
;; COUNTRY: USA
;; ZIP: 19426
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Macintosh
;; OPERATING SYSTEM: System 7.1
;; SOFTWARE: Word 5.1 (Patentin)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/984,186
;; FILING DATE: 29-Oct-2001
;; CLASSIFICATION: <unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/797,689
;; FILING DATE: 31-JAN-1997
;; APPLICATION NUMBER: US 08/256,927
;; FILING DATE: 28-JUL-1994
;; APPLICATION NUMBER: FR 92/01064
;; FILING DATE: 31-JAN-1992
;; APPLICATION NUMBER: PCT/FR93/00085
;; FILING DATE: 28-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith Ph.D., Julie K.
;; REGISTRATION NUMBER: P-38,619
;; REFERENCE/DOCKET NUMBER: ST92006-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (610) 454-3839
;; TELEFAX: (610) 454-3808
;; INFORMATION FOR SEQ ID NO: 2:

```
/
/ SEQUENCE CHARACTERISTICS:
/   LENGTH: 610 amino acids
/   TYPE: amino acid
/   TOPOLOGY: linear
/   MOLECULE TYPE: protein
/   SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match      100.0%; Score 70; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 5.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      100 TVATLRETYGEMAD 113

RESULT 15
US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
;   APPLICANT: Bell et al.
;   TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
;   FILE REFERENCE: PF556
;   CURRENT APPLICATION NUMBER: US/10/153,064
;   CURRENT FILING DATE: 2002-05-24
;   PRIOR APPLICATION NUMBER: 60/293,212
;   PRIOR FILING DATE: 2001-05-25
;   NUMBER OF SEQ ID NOS: 137
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 133
;   LENGTH: 651
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-153-064-133

Query Match      100.0%; Score 70; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 6.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TVATLRETYGEMAD 14
Db      142 TVATLRETYGEMAD 155

Search completed: April 19, 2004, 12:05:19
Job time : 2.14589 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.09695 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_92_100

Perfect score: 47

Sequence: 1 AKQEPERNE 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	600	2 A47391	serum albumin prec
2	47	100.0	608	1 ABRTS	serum albumin prec
3	47	100.0	609	1 ABHUS	serum albumin prec
4	47	100.0	609	2 JC5838	albumin - Mongolia
5	43	91.5	453	2 A05139	serum albumin - mo
6	43	91.5	607	1 ABBOS	serum albumin prec
7	43	91.5	607	1 ABBOS	serum albumin prec
8	43	91.5	607	1 ABBOS	serum albumin prec
9	40	85.1	605	1 ABFGS	serum albumin prec
10	39	83.0	608	2 S57632	serum albumin prec
11	35	74.5	614	2 S59517	serum albumin prec
12	35	74.5	615	1 ABBOS	serum albumin prec
13	35	74.5	884	2 G36787	serum albumin prec
14	35	74.5	1668	2 A02722	protein T4012.9 [i
15	34	72.3	377	1 B04358	IgA-specific metal
16	34	72.3	377	1 B04358	varB10 protein - A
17	34	72.3	680	2 A43800	component of type
18	34	72.3	787	2 A48819	nuclear autoantige
19	33	70.2	245	2 C55206	nuclear autoantige
20	33	70.2	245	2 C51243	thiamin biosynthes
21	33	70.2	245	2 A36091	thiamin biosynthes
22	33	70.2	343	2 H84108	hypothetical prote
23	33	70.2	399	2 H84108	ABC transporter BH
24	33	70.2	407	2 G83477	hypothetical sigma-70
25	33	70.2	433	2 B82965	hypothetical prote
26	33	70.2	489	2 F70401	Flagellar M-ring p
27	33	70.2	607	1 A3XL72	74K albumin precu
28	33	70.2	713	2 AC3220	catalase C [import
29	33	70.2	760	2 JQ1383	hypothetical 84K p

RESULT 1

A47391

serum albumin precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)

C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999

C:Accession: A47391

R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993

A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilin

A:Reference number: A47391; MUID:93211971; PMID:8460152

A:Contents: B/B homozygote

A:Accession: A47391

A>Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-600 <WAT>

A:Cross-references: GB:M90463; NID:g342294; PIDN:AAA36906.1; PID:g342295

A:Experimental source: liver

A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)

C:Superfamily: serum albumin; serum albumin repeat homology

F;21-194/Domain: serum albumin repeat homology <SA1>

F;213-386/Domain: serum albumin repeat homology <SA2>

F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 47; DB 2; Length 600;

Best Local Similarity 100.0%; Pred. No. 0.26;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9

DB 108 AKQEPERNE 116

RESULT 2

ABRTS

serum albumin precursor - rat

N;Alternate names: preproalbumin

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 22-Jun-1999

C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233

R:Sargent, T.D.; Yang, M.; Bonner, J.

Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981

A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.

A:Reference number: A93872; MUID:81223722; PMID:7017712

A:Accession: A93872

A:Molecule type: mRNA

A:Residues: 1-608 <SAR>

A:Cross-references: GB:V01222; GS:J00698; NID:G55627; PIDN:CAA24532.1; PID:g55628

R:Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.

J. Biol. Chem. 252, 6846-6855, 1977

A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analy:

A:Reference number: A92211; MUID:77249657; PMID:893447

A>Note: cleavages during protein maturation

30 33 70.2 906 2 G96621 probable disease r
31 33 70.2 1364 2 T40839 hypothetical prote
32 32 68.1 106 2 T00722 hypothetical prote
33 32 68.1 175 2 T02557 hypothetical prote
34 32 68.1 207 2 T10285 hypothetical prote
35 32 68.1 412 2 F91008 probable oxidoredu
36 32 68.1 412 2 H85852 probable oxidoredu
37 32 68.1 486 2 G94433 hypothetical prote
38 32 68.1 559 2 T12680 peroxisomal target
39 32 68.1 721 2 D82934 ATP-dependent zinc
40 32 68.1 732 2 S23001 tral protein - Esc
41 32 68.1 1032 2 G89427 protein T08D2.3 [i
42 32 68.1 1156 2 T43326 germline RNA helic
43 32 68.1 1172 2 T32759 hypothetical prote
44 32 68.1 1291 2 T00019 period protein hom
45 32 68.1 3051 2 S42373 hypothetical prote

ALIGNMENTS

A/Accession: A92211
A/Molecule type: protein
A/Residues: 1-38 <STR>
R/Isomura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A/Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A/Reference number: A91946; MUID:78109429; PMID:564345
A/Accession: A91946
A/Molecule type: protein
A/Residues: 25-222 <IS1>
R/Isomura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A/Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino acid sequence of the N-terminal fragment
A/Reference number: A91940; MUID:76260153; PMID:956149
A/Accession: A91940
A/Molecule type: protein
A/Residues: 223-288/572-608 <IS2>
A/Note: 262-Leu was also found
R/Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A/Title: Copper(II)-binding ability of human alpha-fetoprotein.
A/Reference number: A90758; MUID:79001617; PMID:80265
A/Contents: annotation; copper binding
R/Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A/Title: Structures of histamine-releasing peptides formed by the action of acid protease
A/Reference number: A45800; MUID:89341406; PMID:2474609
A/Accession: C45800
A/Status: preliminary
A/Molecule type: protein
A/Residues: 166-173 <AR>
R/Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A/Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved method
A/Reference number: 157621; MUID:87286876; PMID:3475566
A/Accession: 157621
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-5 <RES>
A/Cross-references: GB:M16825; MUID:9202828; PIDN:AAA40712.1; PID:9554412
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status experimental <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-608/Product: serum albumin #status experimental <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper [His] #status experimental
F:77-86, 99-115, 114-125, 148-193, 192-201, 224-270, 269-277, 289-303, 302-313, 340-385, 384-393, 4

Query Match 100.0%; Score 47; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
Db 116 AKQEPERNE 124
|||||

RESULT 3
ABUS
serum albumin precursor [validated] - human
N/Alternate names: preproalbumin
N/Contains: kinetensin
C/Species: Homo sapiens (man)
C/Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000
C/Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S36
R/Law, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur
Nucleic Acids Res. 9, 6103-6114, 1981
A/Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli
A/Reference number: A93743; MUID:82081882; PMID:6171778
A/Accession: A93743

A/Molecule type: mRNA
A/Residues: 1-419, 'K', 421-609 <LAW>
A/Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:928591; PIDN:CAA2.
R/Dugaiczyk, A.; Law, S.W.; Dennison, C.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A/Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A/Reference number: A93936; MUID:82105994; PMID:6275391
A/Accession: A93936
A/Molecule type: mRNA
A/Residues: 1-120, 'G', 122-609 <DUG>
A/Cross-references: EMBL:V00494; NID:928589; PIDN:CAA23753.1; PID:928590
R/Urano, Y.; Watanabe, K.; Sakai, M.; Tamaki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A/Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A/Reference number: I39427; MUID:86140099; PMID:2419329
A/Accession: I39427
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-26 <URA>
A/Cross-references: GB:M13075; NID:9178330; PIDN:AAA51688.1; PID:9553173
R/Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A/Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family
A/Reference number: I59286; MUID:94181575; PMID:8134387
A/Accession: I59286
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 282-290, 'KSRFDLQ' <WAT>
A/Cross-references: GB:S69192; NID:9546032; PIDN:AA30282.1; PID:9546033
A/Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R/Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A/Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl
A/Reference number: I59313; MUID:94294404; PMID:8022807
A/Accession: I59313
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 589-590, 'ALPRVKNLLQVKLP' <MAD>
A/Cross-references: GB:S70799; NID:9547231; PIDN:AA31177.1; PID:9547232
A/Note: this frame-shift variant is designated albumin Bazzano; four additional variant
R/Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A/Reference number: G08292
A/Accession: G01747
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-120, 'G', 122-455 <MEN>
A/Cross-references: EMBL:U22961; NID:9763428; PIDN:AAA64922.1; PID:9763431
R/Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A/Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex
A/Reference number: S55314; MUID:95275251; PMID:7755581
A/Accession: S55314
A/Molecule type: protein
A/Residues: 19-27 <LED>
R/Meloun, B.; Moravsek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A/Title: Complete amino acid sequence of human serum albumin.
A/Reference number: A91420; MUID:76187907; PMID:1225573
A/Accession: A91420
A/Molecule type: protein
A/Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395
R/Roehr, U.; Spittler, G.; Trippier, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A/Title: Isolation and structure elucidation of middle-molecular weight peptides from u
A/Reference number: S06422
A/Note: this paper is in German, with an English abstract
A/Accession: S06422
A/Molecule type: protein
A/Residues: 25-48 <ROE>
R/Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A/Title: Mass spectrometric identification of modifications to human serum albumin t

A:Reference number: S36882; MUID:93384321; PMID:8373198
 A:Accession: S36882
 A:Molecule type: protein
 A:Residues: 45-67;141-160;311-337;469-490;570-581 <FIN>
 R:Kausler, E.; Spiteiler, G.
 Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
 A:Title: Bruchatuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol
 A:Reference number: S14999; MUID:92126241; PMID:11772598
 A:Accession: S17599
 A:Molecule type: protein
 A:Residues: 25-54;354-357;431-447 <KAU>
 A:Note: 49-Leu was also found
 R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A:Title: Structures of bisamine-releasing peptides formed by the action of acid protease
 A:Reference number: A45800; MUID:89341406; PMID:2474609
 A:Accession: A45800
 A:Molecule type: protein
 A:Residues: 166-173 <CAR>
 R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
 Biochem. Biophys. Res. Commun. 136, 983-988, 1986
 A:Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre
 A:Reference number: A03239; MUID:86242180; PMID:3087352
 A:Accession: A03239
 A:Molecule type: protein
 A:Residues: 166-173, 'L' <MOG>
 R:Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
 Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
 A:Title: Mutations in genetic variants of human serum albumin found in Italy.
 A:Reference number: A38255; MUID:91062352; PMID:2247440
 A:Accession: C38255
 A:Molecule type: protein
 A:Residues: 76-111 <GAL1>
 A:Accession: B38255
 A:Molecule type: protein
 A:Residues: 82-105, 'K', 107-110 <GAL2>
 A:Note: this variant is designated albumin Vibo Valentia
 A:Accession: A38255
 A:Molecule type: protein
 A:Residues: 76-83, 'K', 85-106 <GAL3>
 A:Note: this variant is designated albumin Torino
 R:Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
 Eur. J. Biochem. 214, 437-444, 1993
 A:Title: The structural characterization and bilirubin-binding properties of albumin Her
 A:Reference number: S33298; MUID:9252504; PMID:8513793
 A:Accession: S33298
 A:Molecule type: protein
 A:Residues: 255-263, 'E', 265-281 <MIN1>
 A:Note: this variant is designated albumin Herborn
 R:Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,
 Biochim. Biophys. Acta 1119, 232-238, 1992
 A:Title: Two albumins with identical electrophoretic mobility are produced by differ
 A:Reference number: S21078; MUID:92190239; PMID:1347703
 A:Accession: S21078
 A:Molecule type: protein
 A:Residues: 354-356, 'K', 358-378 <MIN2>
 A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,
 R:He, X.M.; Carter, D.C.
 Nature 358, 209-215, 1992
 A:Title: Atomic structure and chemistry of human serum albumin.
 A:Reference number: A46756; MUID:92334427; PMID:1630489
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms
 R:Brown, J.R.; Shockley, P.; Behrens, P.Q.
 In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,
 A:Reference number: A94442
 A:Contents: annotation; three-dimensional structure and disulfide bonds
 R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.
 Collect. Czech. Chem. Commun. 42, 564-579, 1977
 A:Title: Disulfide bonds in human serum albumin.
 A:Reference number: A9030
 A:Contents: annotation; disulfide bonds
 R:Jacobsen, C.
 Biochem. J. 171, 453-459, 1978

A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
 A:Reference number: A90299; MUID:78186630; PMID:656055
 A:Contents: annotation; bilirubin-binding site
 R:Peters, T.; Reed, R.G.
 in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20
 A:Title: Serum albumin: conformation and active sites.
 A:Reference number: A34408
 A:Contents: annotation; binding sites
 R:Harper, M.E.; Dugalczyk, A.
 Am. J. Hum. Genet. 35, 565-572, 1983
 A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene:
 A:Reference number: A90028; MUID:83279982; PMID:6192711
 A:Contents: annotation; gene position
 R:Walker, J.E.
 FEBS Lett. 66, 173-175, 1976
 A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid
 A:Reference number: A46755; MUID:76257808; PMID:955075
 A:Contents: annotation
 A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic aci
 R:Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
 FEBS Lett. 298, 265-268, 1992
 A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phospho
 A:Reference number: A56294; MUID:92183881; PMID:1544460
 A:Contents: annotation
 A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in 1
 atase activity
 C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
 C:Comment: A large number of variants of human serum albumin have been described.
 C:Genetics:
 A:Gene: GDB:ALB
 A:Cross-references: GDB:118990; OMIM:103600
 A:Map position: 4q11-q13
 A:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox
 F19-24/Domain: signal sequence #status predicted <SIG>
 F125-609/Product: serum albumin #status experimental <PRO>
 F29-202/Domain: serum albumin repeat homology <SA1>
 F166-174/Product: kinetensin #status experimental <KIP>
 F221-394/Domain: serum albumin repeat homology <SA2>
 F413-592/Domain: serum albumin repeat homology <SA3>
 F27/Binding site: copper (His) #status predicted
 F177-86-99-115-114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,
 F214/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental
 Query Match 100.0%; Score 47; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKQEPERNE 9
 DB 116 AKQEPERNE 124
 RESULT 4
 JC5838
 albumin - Mongolian jird
 C:Species: Meriones unguiculatus (Mongolian jird)
 C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000
 R:Yoshida, K.; Seto-Ohashima, A.; Sinohara, H.
 DNA Res. 4, 351-354, 1997
 A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the
 A:Reference number: JC5838; MUID:98116663; PMID:9455485
 A:Accession: JC5838
 A:Molecule type: mRNA
 A:Residues: 1-609 <YOS>
 A:Cross-references: DDBJ:AB006197; NID:G2317277; PIDN:BAA21765.1; PID:G2317278
 A:Experimental source: liver
 C:Superfamily: serum albumin; serum albumin repeat homology
 F:222-395/Domain: serum albumin repeat homology <SA2>

Query Match 100.0%; Score 47; DB 2; Length 609;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
 |||||
 DB 117 AKQEPERNE 125

RESULT 5
 A05139
 serum albumin - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
 C:Accession: A05139; 148638
 R:Minghetti, P.P.; Law, S.W.; Dugaiczky, A.
 Mol. Biol. Evol. 2, 347-358, 1985
 A>Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudogenes
 A:Reference number: A93055; MUID:88216123; PMID:2452956
 A:Accession: A05139
 A:Molecule type: mRNA
 A:Residues: 1-418 <MIN>
 A:Cross-references: GB:ML6111; NID:9191764; PIDN:AAA37190.1; PID:G191765
 R:Boccaccio, C.; Deschaitrette, J.; Meunier-Rotival, M.
 Gene 88, 181-186, 1990
 A>Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the genome
 A:Reference number: 148638; MUID:90269606; PMID:1971802
 A:Accession: 148638
 A>Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 379-453 <BOC>
 A:Cross-references: EMBL:X13060; NID:952939; PIDN:CAA1459.1; PID:9899334
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>
 F:123-286/Domain: serum albumin repeat homology <SA2>
 F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match 91.5%; Score 43; DB 2; Length 453;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KQEPERNE 9
 |||||
 DB 19 KQEPERNE 26

RESULT 6
 ABBOS
 serum albumin precursor [validated] - bovine
 N:Alternate names: 67K protein; prealbumin
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
 C:Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94551
 R:Holowachuk, E.W.; Stoltzberg, J.K.; Reed, R.G.; Peters Jr., T.
 Submitted to the EMBL Data Library, August 1991
 A:Description: Bovine serum albumin: cDNA sequence and expression.
 A:Reference number: A38885
 A:Accession: A38885
 A:Molecule type: mRNA
 A:Residues: 1-607 <HOL>
 R:Hirayana, K.; Akashi, S.; Furiya, M.; Fukuhara, K.
 Biochem. Biophys. Res. Commun. 173, 639-646, 1990
 A>Title: Rapid confirmation and revision of the primary structure of bovine serum albumin
 A:Reference number: A36401; MUID:91083649; PMID:2260975
 A:Accession: A36401
 A:Molecule type: protein
 A:Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>
 R:MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
 Eur. J. Biochem. 98, 477-485, 1979
 A>Title: Biosynthesis of bovine plasma proteins in a cell-free system.
 A:Reference number: A91258; MUID:80024278; PMID:488109

A:Accession: A91258
 A:Molecule type: protein
 A:Residues: 1-32 <WAG>
 R:Hsieh, J.C.; Lin, F.P.; Tam, M.F.
 Anal. Biochem. 170, 1-8, 1988
 A>Title: Electrophoretic transfer of an analytical isoelectrofocusing
 A:Reference number: A60808; MUID:88267456; PMID:3389500
 A:Accession: B60808
 A:Molecule type: protein
 A:Residues: 25-41 <HSI>
 R:Strawich, E.; Glimcher, M.J.
 Eur. J. Biochem. 191, 47-56, 1990
 A>Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelins' is alb
 A:Reference number: S10780; MUID:90336641; PMID:2379503
 A:Accession: S10780
 A:Molecule type: protein
 A:Residues: 25-41, 'H', 43-57, 59-64 <STR>
 R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A>Title: Structures of histamine-releasing peptides formed by the action of acid protea
 A:Reference number: A45800; MUID:89341406; PMID:2474609
 A:Accession: D45800
 A:Molecule type: protein
 A:Residues: 163-172 <CAR>
 R:Carraway, R.E.; Mitra, S.P.; Cochran, D.E.
 J. Biol. Chem. 262, 5968-5973, 1987
 A>Title: Structure of a biologically active neurotensin-related peptide obtained from p
 A:Reference number: A26693; MUID:87194805; PMID:2437111
 A:Accession: A26693
 A:Molecule type: protein
 A:Residues: 165-172, 'L', 'CA2'
 R:Reed, R.G.; Putnam, F.W.; Peters Jr., T.
 Biochem. J. 191, 867-868, 1980
 A>Title: Sequence of residues 400-403 of bovine serum albumin.
 A:Reference number: A90309; MUID:82023364; PMID:7283978
 A:Accession: A90309
 A:Molecule type: protein
 A:Residues: 402-433 <REE>
 R:Brown, J.R.
 Fed. Proc. 34, 591, 1975
 A>Title: Structure of bovine serum albumin.
 A:Reference number: A91458
 A:Accession: A91458
 A:Molecule type: protein
 A:Residues: 25-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'I'
 R:Brown, J.R.
 submitted to the Atlas, April 1975
 A:Reference number: A94551
 A:Accession: A94551
 A:Molecule type: protein
 A:Residues: 190-195 <BR2>
 R:Brown, J.R.
 Fed. Proc. 33, 1389, 1974
 A:Reference number: A91457
 A:Contents: annotation; disulfide bonds
 R:Werlen, R.C.; Offord, R.E.; Rose, K.
 Biochem. J. 302, 907-911, 1994
 A>Title: Preparation and characterization of novel substrates of insulin proteinase (EC
 A:Reference number: S55232; MUID:95031935; PMID:7945219
 A:Accession: S55232
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 529-536; 569-572 <WER>
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; copper binding; duplication; plasma
 F:1-18/Domain: signal sequence #status experimental <SIG>
 F:11-24/Domain: propeptide #status experimental <PRO>
 F:23-607/Product: serum albumin #status experimental <MPT>
 F:29-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, 'I'

Query Match 91.5%; Score 43; DB 1; Length 607;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9
 Db 117 KOEPERNE 124
 |||||

RESULT 7
 ABSHS
 serum albumin precursor - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S06936
 R:Brown, W.M.; Dziugielewska, K.M.; Foreman, R.C.; Saunders, N.R.
 Nucleic Acids Res. 17, 10495, 1989
 A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
 A:Reference number: S06936; MUID:9009888; PMID:2602160
 A:Accession: S06936
 A:Molecule type: mRNA
 A:Residues: 1-607 <BRO>
 A:Cross-references: EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387
 C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRO>
 F:25-607/Product: serum albumin #status predicted <MAT>
 F:29-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
 F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 91.5%; Score 43; DB 1; Length 607;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9
 Db 117 KOEPERNE 124
 |||||

RESULT 8
 ABSHS
 serum albumin precursor - horse
 C:Species: Equus caballus (domestic horse)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S34053
 R:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
 Eur. J. Biochem. 215, 205-212, 1993
 A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
 A:Reference number: S34053; MUID:93345495; PMID:8344282
 A:Accession: S34053
 A:Molecule type: mRNA
 A:Residues: 1-607 <HOA>
 A:Cross-references: GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672
 C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRO>
 F:25-607/Product: serum albumin #status predicted <MAT>
 F:29-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4

F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 91.5%; Score 43; DB 1; Length 607;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9
 Db 117 KOEPERNE 124
 |||||

RESULT 9
 ABPGS
 serum albumin precursor - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S01382; A61006
 R>Weinstock, J.; Baldwin, G.S.
 Nucleic Acids Res. 16, 9045, 1988
 A:Title: Nucleotide sequence of porcine liver albumin.
 A:Reference number: S01382; MUID:8901582; PMID:3174440
 A:Accession: S01382
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-605 <WEI>
 A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30370.1; PID:G833798
 R>Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
 J. Bone Miner. Res. 4, 235-241, 1989
 A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral
 A:Reference number: A61006; MUID:89269769; PMID:2728927
 A:Accession: A61006
 A:Molecule type: protein
 A:Residues: 23-51, 'X', 53-54; 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XV', 155 <LIM>
 A:Experimental source: dental enamel
 A:Note: albumin and other serum proteins are also found in bone
 C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
 F:17-22/Domain: propeptide #status predicted <PRO>
 F:23-605/Product: serum albumin #status predicted <MAT>
 F:27-199/Domain: serum albumin repeat homology <SA1>
 F:218-391/Domain: serum albumin repeat homology <SA2>
 F:410-589/Domain: serum albumin repeat homology <SA3>
 F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4
 F:261/Binding site: bilirubin (Lys) #status predicted

Query Match 85.1%; Score 40; DB 1; Length 605;
 Best Local Similarity 87.5%; Pred. No. 6.2;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9
 Db 115 KOEPERNE 122
 |||||

RESULT 10
 S57632
 serum albumin precursor - cat
 C:Species: Felis silvestris catus (domestic cat)
 C>Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
 C:Accession: JC4660; S57632
 R>Hilger, C.; Grigori, F.; Hentges, F.
 Gene 169, 295-296, 1996
 A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
 A:Reference number: JC4660; MUID:96194824; PMID:8647469
 A:Accession: JC4660
 A:Molecule type: mRNA
 A:Residues: 1-608 <HI2>
 A:Cross-references: EMBL:X84842; NID:G986484; PIDN:CAA59279.1; PID:G986485
 A:Experimental source: liver
 C:Comment: This protein is the major protein component in plasma. It functions as a mul

ein has 35 conserved cysteine residues.

C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: liver; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRP>
F:25-608/Product: serum albumin #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match 83.0%; Score 39; DB 2; Length 608;
Best Local Similarity 87.5%; Pred. No. 9.9;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KQEPERNE 9
DB 117 KKEPERNE 124
- - - - -
- - - - -
- - - - -

RESULT 11
S59517
serum albumin precursor - monoclod cobra
C:Species: Naja naja kaouthia, Naja naja siamensis (monoclod cobra)
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 17-Mar-2000
C:Accession: S59517
R:Wang, X.; Havsteen, B.; Hansen, H.
Biol. Chem. Hoppe-Seyler 376, 543-553, 1995
A:Title: Evidence of the coevolution of a snake toxin and its endogenous antitoxin. Clon
A:Reference number: S59517; MUID:96145734; PMID:8561913
A:Note: the source is designated as Naja naja kaouthia, but referred to as Chinese cobra
A:Accession: S59517
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-614 <WAN>
A:Cross-references: EMBL:X78598; NID:g469860; PIDN:CAA55333.1; PID:g469861
C:Superfamily: serum albumin; serum albumin repeat homology
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-614/Product: serum albumin #status predicted <MAT>
F:32-206/Domain: serum albumin repeat homology <SA1>
F:225-398/Domain: serum albumin repeat homology <SA2>
F:417-596/Domain: serum albumin repeat homology <SA3>

Query Match 74.5%; Score 35; DB 2; Length 614;
Best Local Similarity 66.7%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
DB 119 AKADPERNE 127
- - - - -
- - - - -
- - - - -

RESULT 12
ABCHS
serum albumin precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S15571; A05078; A13451
R:Cassady, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A:Reference number: S15571
A:Accession: S15571
A:Molecule type: mRNA
A:Residues: 1-615 <CAS>
A:Cross-references: EMBL:X60688; NID:g63747; PIDN:CAA43098.1; PID:g63748
R:Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A:Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein
A:Reference number: A05078; MUID:83161037; PMID:6187737
A:Accession: A05078
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-28 <HAC>
A:Cross-references: GB:V00381; NID:g63038; PIDN:CAA23680.1; PID:g63039

A;Title: Analysis of the immunoglobulin A protease gene of *Streptococcus sanguis*.
A;Reference number: A60272; PMID:91100011; PMID:1997065
A;Accession: A60272
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-1668 <GIL>
C;Comment: This protein is shown from the start of translation of this gene as determined the start codon is shown in entry B60272.
C;Genetics:
A;Gene: iga
A;Start codon: GTG
C;Superfamily: *Streptococcus sanguis* Iga-specific metalloendopeptidase
C;Keywords: hydrolase; metalloproteinase; tandem repeat; zinc

Query Match 74.5%; Score 35; DB 2; Length 1668;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KOEPEKN 8
Db 353 KOEPEKN 359

RESULT 15
BOAG58
virB10 protein - *Agrobacterium tumefaciens* plasmid pTiC58
C;Species: *Agrobacterium tumefaciens*
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C;Accession: S12350; S11835; S10525
R;Shirasu, K.; Morel, P.; Kado, C.I.
Mol. Microbiol. 4, 1153-1163, 1990
A;Title: Characterization of the virB operon of an *Agrobacterium tumefaciens* Ti plasmid.
A;Reference number: S12341; PMID:91041724; PMID:2233252
A;Accession: S12350
A;Molecule type: DNA
A;Residues: 1-377 <SHI>
A;Cross-references: EMBL:J03320; NID:gl54781; PIDN:AAA91600.1; PID:gl54792
R;Rogowsky, P.M.; Powell, B.S.; Shirasu, K.; Lin, T.S.; Morel, P.; Zyprian, E.M.; Steck, Plasmid 23, 85-106, 1990
A;Title: Molecular characterization of the vir regulon of *Agrobacterium tumefaciens*: com
A;Reference number: S11825; PMID:90301800; PMID:2194232
A;Accession: S11835
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-377 <ROG>
A;Cross-references: EMBL:J03320; NID:gl54781; PIDN:AAA91600.1; PID:gl54792
R;Kulda, G.A.; de Vos, G.; Owen, G.; McCaffrey, G.; Zambryski, P.
Mol. Gen. Genet. 221, 256-266, 1990
A;Title: The virB operon of *Agrobacterium tumefaciens* pTiC58 encodes 11 open reading fra
A;Reference number: S10515; PMID:90318324; PMID:2370849
A;Accession: S10525
A;Molecule type: DNA
A;Residues: 1-296, A', 298-377 <KUL>
A;Cross-references: EMBL:X53264; NID:g39152; PIDN:CAA37363.1; PID:g39162
C;Genetics:
A;Gene: virB10
A;Genome: plasmid
C;Superfamily: tumor-inducing plasmid pTiC58 virB10 protein

Query Match 72.3%; Score 34; DB 1; Length 377;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AKOEPEKNE 9
Db 99 ARSEPERHE 107

Search completed: April 19, 2004, 12:02:24
Job time : 2.09695 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.65651 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_92_100

Perfect score: 47

Sequence: 1 AKQEPERNE 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	600	1	ALBU_MACMU
2	47	100.0	608	1	ALBU_RAT
3	47	100.0	609	1	ALBU_HUMAN
4	47	100.0	609	1	ALBU_MERIN
5	43	91.5	607	1	ALBU_BOVIN
6	43	91.5	607	1	ALBU_HORSE
7	43	91.5	607	1	ALBU_SHEEP
8	43	91.5	608	1	ALBU_MOUSE
9	40	85.1	605	1	ALBU_PIG
10	40	85.1	608	1	ALBU_CANFA
11	39	83.0	608	1	ALBU_FELCA
12	39	83.0	608	1	ALBU_RABIT
13	35	74.5	501	1	A37C_DROSI
14	35	74.5	615	1	ALBU_CHICK
15	34	72.3	341	1	SUR6_XENLA
16	34	72.3	377	1	VIBX_AGR75
17	34	72.3	680	1	NASP_RASIT
18	34	72.3	788	1	NASP_HUMAN
19	33	70.2	251	1	THIF_ECOLI
20	33	70.2	489	1	FLIF_AQUAE
21	33	70.2	607	1	ALB2_XENLA
22	33	70.2	906	1	RDL4_ARATH
23	32	68.1	207	1	VE18_NPVOP
24	32	68.1	215	1	PS3A_THERAC
25	32	68.1	412	1	YEIT_ECOS7
26	32	68.1	669	1	STM_ENTFA
27	32	68.1	732	1	TRIG_ECOLI
28	32	68.1	733	1	ACE1_TRIRE
29	32	68.1	950	1	STL2_MOUSE
30	32	68.1	1156	1	GLH4_CABEL
31	32	68.1	1291	1	PER1_MOUSE
32	32	68.1	3133	1	HNCT_BOMBO
33	32	68.1	3767	1	MUA3_CABEL

ALIGNMENTS

RESULT 1

ID	ALBU_MACMU	STANDARD;	PRT;	600 AA.
AC	Q28522;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Serum albumin precursor (Fragment).			
GN	ALB.			
OS	Macaca mulatta (Rhesus macaque).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;			
OC	Cercopitheidae; Macaca.			
OX	NCBI_TaxID=9544;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93211971; PubMed=8460152;			
RA	Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,			
RA	Dwulet J., Putnam F.W.			
RT	"CDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding."			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).			
CC	!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.			
CC	!- SUBCELLULAR LOCATION: Secreted.			
CC	!- TISSUE SPECIFICITY: Plasma.			
CC	!- SIMILARITY: Belongs to the ALB/APP/VDB family.			
CC	!- SIMILARITY: Contains 3 albumin domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/ or send an email to license@isb-sib.ch).			
CC	EMBL; M90463; AAA36906.1; -			
DR	PIR; A47391; A47391.			
DR	HSSP; P02768; IE7B.			
DR	InterPro; IPR000264; Serum albumin.			
DR	Pfam; PF00273; transport prot; 3.			
DR	PRINTS; PD00802; SERUMALBUMIN.			
DR	ProDom; PD002486; Serum albumin; 1.			
DR	SMART; SM00103; ALBUMIN_3.			
DR	PROSITE; PS00212; ALBUMIN; 3.			
KW	Metal-binding; Lipid-binding; Repeat; Signal; Copper.			
FT	NON TER	1		
FT	SIGNAL	<1	10	BY SIMILARITY.
FT	PROPEP	11	16	BY SIMILARITY.
FT	CHAIN	17	600	SERUM ALBUMIN.
FT	DOMAIN	17	197	ALBUMIN 1.
FT	DOMAIN	204	389	ALBUMIN 2.
FT	DOMAIN	396	587	ALBUMIN 3.

34	31	66.0	143	1	PSAG_HORVU	Q00327 hordeum vul
35	31	66.0	149	1	RL9_FUSNN	Q8A110 fusobacteri
36	31	66.0	167	1	PSAG_SPIOL	P12357 spinacia ol
37	31	66.0	193	1	RHOC_HUMAN	P08134 homo sapien
38	31	66.0	193	1	RHOC_MOUSE	Q62159 mus musculu
39	31	66.0	310	1	GTRB_BPP22	P57022 bacterioph
40	31	66.0	322	1	RLA0_LEICH	P39096 leishmania
41	31	66.0	323	1	RLA0_LEIN	P39097 leishmania
42	31	66.0	389	1	VPD_BP186	P21679 bacterioph
43	31	66.0	422	1	K3LI_RAT	P43556 rattus norv
44	31	66.0	467	1	MMUP_ECOLI	Q47689 escherichia
45	31	66.0	504	1	A37C_DROME	P19487 drosophila


```

FT METAL 19 19 COPPER (BY SIMILARITY).
FT BINDING 256 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 100.0%; Score 47; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
Db 108 AKQEPERNE 116

RESULT 2
ALBU RAT
ID ALBU RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN ALB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=8123722; PubMed=7017712;
RX Sargent T.D., Yang M., Bonner J.;
RA "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RT Proc. Natl. Acad. Sci. U.S.A. 78:243-246 (1981).
RN [2]
RP SEQUENCE OF 1-38, AND PROCESSING.
RX MEDLINE=77249657; PubMed=893447;
RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
RT piece. Analysis of the direct translation product of albumin
RT messenger RNA.";
RL J. Biol. Chem. 252:6846-6855 (1977).
RN [3]
RP SEQUENCE OF 25-222.
RX MEDLINE=78109429; PubMed=564345;
RA Isemura S., Ikenaka T.;
RT "Amino acid sequences of fragments I and II obtained by cyanogen
RT bromide cleavage of rat serum albumin.";
RL J. Biochem. 83:35-48 (1978).
RN [4]
RP SEQUENCE OF 223-288 AND 572-608.
RX MEDLINE=76260153; PubMed=956149;
RA Isemura S., Ikenaka T.;
RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
RT the amino acid sequences of four fragments.";
RL J. Biochem. 79:1183-1196 (1976).
RN [5]
RP SEQUENCE OF 166-174.

```

```

RC TISSUE=Plasma;
RX MEDLINE=87194805; PubMed=2437111;
RA Carraway R.E., Mitra S.P., Cochran D.E.;
RT "Structure of a biologically active neurotensin-related peptide
RT obtained from pepsin-treated albumin(s).";
RL J. Biol. Chem. 262:5968-5973 (1987).
RN [6]
RP COPPER-BINDING.
RX MEDLINE=79001617; PubMed=80265;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Copper(II)-binding ability of human alpha-fetoprotein.";
CC Cancer Res 38:3483-3486 (1978).
CC -I- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -I- FUNCTION: NRP regulates fat digestion, lipid absorption, and
CC blood flow (Potential).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Plasma.
CC -I- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -I- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01222; CAA24532.1; -
DR PIR; A93872; ABRTS.
DR HSP; P02768; 1E7B.
DR InterPro; IPR000284; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM001103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608 SERUM ALBUMIN.
FT PEPTIDE 166 174 NEUROTENSIN-RELATED PEPTIDE.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT VARIANT 262 262 V -> L.
FT CONFLICT 174 174 Y -> L (IN REF. 5).
SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;

```

```

Query Match 100.0%; Score 47; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AKQEPERNE 9
Db 116 AKQEPERNE 124

RESULT 3

ALBU HUMAN STANDARD; PRT; 609 AA.
AC P02768; Q95574; Q13140; Q9P157; Q9P177; Q9UH53; Q9U0Z0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczky A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX MEDLINE=8201882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczky A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RL Submitted (SEP-1999) to the ENBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver.";
RL Submitted (FEB-1999) to the ENBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
RA Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin.";
RL Submitted (AUG-2002) to the ENBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bustow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RP SEQUENCE OF 25-609.
RA Brown J.R., Shockley P., Behrens P.Q.;
RL (In) Bing D.H. (eds.);
RL The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RP SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the ENBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RP SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RP SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994.";
RL Electrophoresis 15:1459-1465(1994).
RN [14]
RP DISULFIDE BONDS.
RA Saber M.A., Stockbauer P., Moravsek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15]
RP BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin.";
RL Biochem. J. 171:453-459(1978).
RN [16]
RP VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
RT domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RP VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=8068523; PubMed=3479777;
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,

RA	Satoh C., Neel J.V.;	ALBU MERUN	STANDARD;	PRT;	609 AA.
RT	"Amino acid substitutions in inherited albumin variants from	ID	ALBU MERUN		
RT	Amerindian and Japanese populations.;"	AC	O35090;		
RL	Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).	DT	15-JUL-1998 (Rel. 36, Created)		
RP	[18]	DT	15-JUL-1998 (Rel. 36, Last sequence update)		
RP	VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.	DT	28-FEB-2003 (Rel. 41, Last annotation update)		
RP	MEDLINE=89345611; PubMed=2762316;	DE	Serum albumin precursor.		
RA	Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,	GN	ALB.		
RA	Neel J.V., Sakurabayashi I., Putnam F.W.;	OS	Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).		
RT	"Point substitutions in Japanese allolalbumins.;"	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
RL	Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;		
RP	[19]	OC	Meriones.		
RP	VARIANTS MANAUS; OSAKA; NAGOVA; FUKUOKA; HONOLULU AND NEW-GUINEA.	OX	NCBI_TaxID=10047;		
RP	MEDLINE=90115852; PubMed=2104980;	RN	(1)		
RA	Arai K., Madison J., Shimizu A., Putnam F.W.;	RP	SEQUENCE FROM N.A.		
RA	"Point substitutions in albumin genetic variants from Asia.;"	RC	STRAIN=MG5 IDR; TISSUE=Liver;		
RL	Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).	RX	MEDLINE=98116663; PubMed=9455485;		
EN	[20]	RA	Yoshida K., Seto-Onshima A., Sinochara H.;		
RP	DESCRIPTION OF VARIANT REDHILL.	RT	"Sequencing of cDNA encoding serum albumin and its extrahepatic		
RA	Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;	RL	synthesis in the Mongolian gerbil, Meriones unguiculatus.;"		
RT	"Albumin Redhill (-1 Arg, 320 Ala--Thr): a glycoprotein variant of	CC	DNA Res. 4:351-354(1997).		
RT	human serum albumin whose precursor has an aberrant signal peptidase	CC	-1- FUNCTION: Serum albumin, the main protein of plasma, has a good		
RT	cleavage site.;"	CC	binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, and		
RL	Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).	CC	hormones, bilirubin and drugs. Its main function is the regulation		
RP	[21]	CC	of the colloidal osmotic pressure of blood.		
RP	VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.	CC	-1- SUBCELLULAR LOCATION: Secreted.		
RX	MEDLINE=91062352; PubMed=2247440;	CC	-1- TISSUE SPECIFICITY: Plasma.		
RA	Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,	CC	-1- SIMILARITY: Belongs to the ALB/APP/VDB family.		
RA	Watkins S., Putnam F.W.;	CC	-1- SIMILARITY: Contains 3 albumin domains.		
RT	"Mutations in genetic variants of human serum albumin found in	CC			
RL	Italy.;"	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
RL	Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
RP	[22]	CC	the European Bioinformatics Institute. There are no restrictions on its		
RP	VARIANT VENEZIA.	CC	use by non-profit institutions as long as its content is in no way		
RA	Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,	CC	modified and this statement is not removed. Usage by and for commercial		
RA	Minchiotti L., Putnam F.W.;	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
RT	"A donor splice mutation and a single-base deletion produce two	CC	or send an email to license@isb-sib.ch).		
RT	carboxyl-terminal variants of human serum albumin.;"	CC			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).	CC			
RP	[23]	DR	EMBL; AB006197; BAA21765.1; -.		
RP	VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;	DR	PIR; JCS838; JCS838.		
RP	KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.	DR	HSSP; P02768; 1E7B.		
RX	MEDLINE=92052189; PubMed=1946412;	DR	InterPro; IPR000264; Serum albumin.		
RA	Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,	DR	Pfam; PF00273; transport prot.; 3.		
RA	Matsuda Y.-I., Anaki I., Putnam F.W.;	DR	PRINTS; PR00802; SERUMALBUMIN.		
RT	"Genetic variants of serum albumin in Americans and Japanese.;"	DR	ProDom; PD002486; Serum albumin; 1.		
RL	Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).	DR	SMART; SM00103; ALBUMIN; 3.		
RP	[24]	DR	PROSITE; PS00212; ALBUMIN; 3.		
RP	VARIANT CASEBROOK ASN-518.	KW	Metal-binding; Lipid-binding; Repeat; Signal; Copper.		
RX	MEDLINE=91316157; PubMed=1859851;	FT	SIGNAL 1 18		
RA	Peach R.J., Brennan S.O.;	FT	PROPEP 19 24		
RT	"Structural characterization of a glycoprotein variant of human serum	FT	CHAIN 25 609		
RT	albumin: albumin Casebrook (494 Asp-->Asn).;"	FT	DOMAIN 25 206		
RL	Biochim. Biophys. Acta 1097:49-54(1991).	FT	DOMAIN 213 398		
RP	[25]	FT	DOMAIN 405 596		
RP	VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.	FT	METAL 28 28		
RX	MEDLINE=92190239; PubMed=1347703;	FT	DISULFID 78 87		
RA	Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,	FT	DISULFID 100 116		
RA	Rochu D., Porta F.;	FT	DISULFID 115 126		
RT	"Two allolalbumins with identical electrophoretic mobility are produced	FT	DISULFID 149 194		
Query Match	100.0%; Score 47; DB 1; Length 609;	FT	DISULFID 193 202		
Best Local Similarity	100.0%; Pred. No. 0.17;	FT	DISULFID 225 271		
Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	FT	DISULFID 270 278		
QY	1 AKOEPERNE 9	FT	DISULFID 290 304		
Db	116 AKOEPERNE 124	FT	DISULFID 303 314		
RESULT 4		FT	DISULFID 341 386		
		FT	DISULFID 385 394		
		FT	DISULFID 417 463		
		FT	DISULFID 462 473		
		FT	DISULFID 486 502		
		FT	DISULFID 501 512		
		FT	DISULFID 539 584		
		FT	DISULFID 583 592		
		SQ	SEQUENCE 609 AA; 9CA5F97F67EF1A48 CRC64;		

```
Query Match      100.0%; Score 47; DB 1; Length 609;
Best Local Similarity 100.0%; Pred No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPNER 9
DB      117 AKQEPNER 125

RESULT 5
ALBU_BOVIN
ID ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Holowachuk E.W., Stoltzenberg J.K., Reed R.G., Peters T. Jr.;
RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RN [3]
RP SEQUENCE FROM N.A.
RA Barry T., Power S., Gannon F.;
RC TISSUE=Liver;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T., Huang M.C.;
RT "The complete cDNA sequence of bovine serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE OF 1-32.
RX MEDLINE=80024278; PubMed=488105;
RA McGilivray R.T.A., Chung D.W., Davie E.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
terminal sequence of prealbumin.";
RL Eur. J. Biochem. 98:477-485(1979).
[6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA Brown J.R.;
RT "Structure of bovine serum albumin.";
RL Fed. Proc. 34:591-591(1975).
[7]
RP REVISIONS TO 190-195.
RA Brown J.R.;
RL Submitted (APR-1975) to the PIR data bank.
[8]
RP SEQUENCE OF 402-433.
RX MEDLINE=82023364; PubMed=7283978;
RA Reed R.G., Putnam F.W., Peters T. Jr.;
RT "Sequence of residues 400-403 of bovine serum albumin.";
RL Biochem. J. 191:867-868(1980).
[9]
RP SEQUENCE OF 19-28.
RX MEDLINE=77134075; PubMed=843354;
RA Patterson J.E., Geller D.M.;
RT "Bovine microsomal albumin: amino terminal sequence of bovine
proalbumin.";
RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
[10]

RX SEQUENCE, AND REVISIONS TO 118-119 AND 180.
RX MEDLINE=91083649; PubMed=2260975;
RA Hirayana K., Akashi S., Furuya M., Fukuhara K.-I.;
RT "Rapid confirmation and revision of the primary structure of bovine
serum albumin by ESIMS and Frit-FAB LC/MS.";
RL Biochem. Biophys. Res. Commun. 173:639-646(1990).
[11]
RP SEQUENCE OF 25-41.
RX MEDLINE=88267456; PubMed=3389500;
RA Hsieh J.C., Lin F.P., Tam M.F.;
RT "Electroblotting onto glass-fiber filter from an analytical
isoelectrofocusing gel: a preparative method for isolating proteins
for N-terminal microsequencing.";
RL Anal. Biochem. 170:1-8(1988).
[12]
RP SEQUENCE OF 437-451.
RA Vilbois F.;
RL Submitted (AUG-1998) to Swiss-Prot.
[13]
RP DISULFIDE BONDS.
RA Brown J.R.;
RT "Structure of serum albumin: disulfide bridges.";
RL Fed. Proc. 33:1389-1389(1974).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; W73993; AAA51411.1; -
EMBL; X58989; CAA41735.1; -
EMBL; Y17769; CAA76847.1; -
EMBL; AF542068; AAN17824.1; -
HSSP; P02768; 1E7B.
InterPro: IPR000264; Serum albumin.
Pfam; PF00273; transport_prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
Polymorphism. 1 18
SIGNAL 19 24 SERUM ALBUMIN.
PROPEP 19 24 ALBUMIN 1.
CHAIN 25 607 ALBUMIN 1.
DOMAIN 25 204 ALBUMIN 2.
DOMAIN 211 396 ALBUMIN 3.
DOMAIN 403 594 COPPER (BY SIMILARITY).
METAL 27 27
DISULFID 77 86
DISULFID 99 115
DISULFID 114 125
DISULFID 147 192
DISULFID 191 200
DISULFID 223 269
DISULFID 268 276
DISULFID 288 302
DISULFID 301 312
DISULFID 339 384
DISULFID 383 392
DISULFID 415 461
```

```
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
FT VARIANT 214 214
FT CONFLICT 302 302
FT CONFLICT 304 305
FT CONFLICT 324 324
FT CONFLICT 394 395
FT CONFLICT 437 437
FT CONFLICT 493 494
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;

Query Match 91.5%; Score 43; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 0.97;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9
Db 117 KOEPERNE 124

RESULT 6
ALBU HORSE
ID ALBU HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)
at 0.27-nm resolution."
RL Eur. J. Biochem. 215:205-212 (1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IGE.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X74045; CAA52194.1; --
CC PIR; S34053; ABHOS.
CC HSSP; P02758; 1E7E.
CC InterPro; IPR000284; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Allergen.
CC SIGNAL 1 18
BY SIMILARITY.
```

```
FT PROPEP 19 24
FT CHAIN 25 607
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 451
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 68598 MW; 258F6E830A1B90C5 CRC64;

Query Match 91.5%; Score 43; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 0.97;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9
Db 117 KOEPERNE 124

RESULT 7
ALBU SHEEP
ID ALBU SHEEP STANDARD; PRT; 607 AA.
AC P14639;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=30038888; PubMed=2602160;
RA Brown W.W., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin."
RL Nucleic Acids Res. 17:10495-10495 (1989).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X17055; CAA34903.1; --
```

DR PIR; S06936; ABSHS.
 DR HSP; P02768; 1E7B.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transprot prot; 3.
 DR PRINTS; PRC0802; SERUMALBUMIN
 DR PRODOM; PRC02486; Serum albumin; 1.
 DR SMART; SMO0103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 607
 FT DOMAIN 25 204
 FT DOMAIN 211 396
 FT DOMAIN 403 594
 FT METAL 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 147 192
 FT DISULFID 191 200
 FT DISULFID 223 269
 FT DISULFID 268 276
 FT DISULFID 288 302
 FT DISULFID 301 312
 FT DISULFID 339 384
 FT DISULFID 393 392
 FT DISULFID 415 461
 FT DISULFID 460 471
 FT DISULFID 484 500
 FT DISULFID 499 510
 FT DISULFID 537 582
 FT DISULFID 581 590
 SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;
 Query Match 91.5%; Score 43; DB 1; Length 607;
 Best Local Similarity 100.0%; Pred. No. 0.97;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 KOEPERNE 9
 Db 117 KOEPERNE 124
 RESULT 8
 ID ALBU MOUSE STANDARD; PRT; 608 AA.
 AC P07724; O61802;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Serum albumin precursor.
 GN ALB OR ALB1 OR ALB-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RA Van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann J., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Liver;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Haney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 99-516 FROM N.A.
 RX MEDLINE=88216123; PubMed=24529556;
 RA Minghetti P.P., Law S.W., Dugaiczky A.;
 RT "The rate of molecular evolution of alpha-fetoprotein approaches that
 of pseudogenes";
 RL Mol. Biol. Evol. 2:347-358(1985).
 RN [5]
 RP SEQUENCE OF 477-551 FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=90269606; PubMed=1971802;
 RA Bocaccio C., Deschatrette J., Meunier-Rotival M.;
 RT "Empty and occupied insertion site of the truncated LINE-1 repeat
 located in the mouse serum albumin-encoding gene.";
 RL Gene 88:181-186(1990).
 RN [6]
 RP SEQUENCE OF 25-44.
 RC TISSUE=Liver;
 RX MEDLINE=93162044; PubMed=1286668;
 RA Giometti C.S., Taylor J., Tollaksen S.L.;
 RT "Mouse liver protein database: a catalog of proteins detected by two-
 dimensional gel electrophoresis.";
 RL Electrophoresis 13:970-991(1992).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AJ011413; CAA09617.1; -
CC EMBL; AK010025; BAB26650.1; -
CC EMBL; BC049971; AAH49971.1; -
CC EMBL; M16111; AAA37190.1; -
CC EMBL; X13060; CAA31458.1; -
CC PIR; A05139; A05139.
CC HSSP; P02768; I27B.
CC SWISS-2DPAGE; P07724; MOUSE.
CC MGD; MGI:87991; Albi.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC PRODOM; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608
FT DOMAIN 25 205
FT DOMAIN 212 297
FT DOMAIN 404 595
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 132 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
FT CONFLICT 27 27
FT CONFLICT 33 33
FT CONFLICT 41 41
SQ SEQUENCE 608 AA; 68692 MW; 292F7C7EBD3A61B4 CRC64;
Query Match 91.5%; Score 43; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 KOEPERNE 9
DB 117 KOEPERNE 124

RC TISSUE=Liver;
RX MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S.; Weinstock J.;
RT "Nucleotide sequence of porcine liver albumin."
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X12422; CAA30970.1; -
CC EMBL; M36787; AAA30988.1; -
CC PIR; S01382; ABPGS.
CC HSSP; P02768; I27B.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC PRODOM; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON_TER 1 1
FT SIGNAL <1 16
FT PROPEP 17 22
FT CHAIN 23 605
FT DOMAIN 23 202
FT DOMAIN 209 394
FT DOMAIN 401 592
FT METAL 31 31
FT DISULFID 75 84
FT DISULFID 97 113
FT DISULFID 112 123
FT DISULFID 145 190
FT DISULFID 189 198
FT DISULFID 221 267
FT DISULFID 266 274
FT DISULFID 286 300
FT DISULFID 299 310
FT DISULFID 337 382
FT DISULFID 381 390
FT DISULFID 413 459
FT DISULFID 458 469
FT DISULFID 482 498
FT DISULFID 497 508
FT DISULFID 535 580
FT DISULFID 579 588
FT CONFLICT 562 562
SQ SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1F4FF CRC64;
Query Match 85.1%; Score 40; DB 1; Length 605;
Best Local Similarity 87.5%; Pred. No. 3.7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 KOEPERNE 9
DB 115 KOEPERNE 122

RESULT 10
ID ALBU_CANFA STANDARD; PRT; 608 AA.
ALBU_CANFA

P49822; Q77705; Q9TS24;
01-OCT-1996 (Rel. 34, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).
GN ALB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RA Hilger C.;
RN Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=20148667; PubMed=10669848;
RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H.,
RA Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog
RT albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285(2000).
RN [3]
RP SEQUENCE OF 25-48.
RX MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W., Sarkar B.;
RT "Isolation, amino acid sequence and copper(II)-binding properties of
RT peptide (1-24) of dog serum albumin.";
RL J. Biol. Chem. 249:5872-5877(1974).
RN [4]
RP SEQUENCE OF 25-38.
RC TISSUE=Heart;
RX MEDLINE=98183340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2PAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
RN [5]
RP SEQUENCE OF 215-478 FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=94201492; PubMed=7512102;
RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,
RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
RT "Molecular characterization of dog albumin as a cross-reactive
RT allergen.";
RL J. Allergy Clin. Immunol. 93:614-627(1994).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Plasma.
CC -1- ALLERGEN: Causes an allergic reaction in human.
CC -1- SIMILARITY: Belongs to the ALB/APP/VBP family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; AJ133489; CAB64867.1; -;
CC EMBL; Y17737; CAAY6841.1; -;
CC EMBL; S72946; AAB30434.1; -;
CC HSP; P02768; 1E7B.
CC HSC-2PAGE; P49822; DOG.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.

PRINTS; P00802; SERUMALBUMIN.
ProDom; P002486; Serum albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
GN ALB.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608
FT DOMAIN 25 205
FT DOMAIN 212 397
FT DOMAIN 404 595
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 289 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
FT CONFLICT 1 26
FT CONFLICT 146 146
FT CONFLICT 206 206
FT CONFLICT 349 349
FT CONFLICT 359 359
FT CONFLICT 448 448
FT CONFLICT 474 474
SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;
Query Match 85.1%; Score 40; DB 1; Length 608;
Best Local Similarity 87.5%; Pred. No. 3.7; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 0;
QY 2 KOEPERNE 9
DB 117 KOEPERNE 124
RESULT 11
ALBU_FELCA STANDARD; PRT; 608 AA.
AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96194824; PubMed=8647469;
RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL Gene 169:295-296(1996).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.

CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X84842; CAA59279.1; -.
CC PIR; JC4660; S57632.
CC HSSP; P02768; S57632.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
CC SIGNAL 1 18
CC PROPEP 19 24
CC CHAIN 25 608
CC DOMAIN 25 205
CC DOMAIN 212 397
CC DOMAIN 404 595
CC METAL 27 27
CC METAL 77 86
CC DISULFID 77 86
CC DISULFID 99 115
CC DISULFID 114 125
CC DISULFID 148 193
CC DISULFID 182 201
CC DISULFID 224 270
CC DISULFID 269 277
CC DISULFID 289 303
CC DISULFID 302 313
CC DISULFID 340 385
CC DISULFID 384 393
CC DISULFID 416 462
CC DISULFID 461 472
CC DISULFID 485 501
CC DISULFID 500 511
CC DISULFID 538 583
CC DISULFID 582 591
CC SEQUENCE 608 AA; 66659 MW; 075629CAC5F60E5F CRC64;
Query Match 83.0%; Score 39; DB 1; Length 608;
Best Local Similarity 87.5%; Pred. No. 5.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 KOEPERNE 9
Db 117 KKEPERNE 124
RESULT 12
ALBU RABIT STANDARD; PRT; 608 AA.
AC P49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;

RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U18344; AAB58347.1; -.
CC HSSP; P02768; L7B.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
CC SIGNAL 1 18
CC PROPEP 19 24
CC CHAIN 25 608
CC DOMAIN 25 205
CC DOMAIN 212 397
CC DOMAIN 404 595
CC METAL 27 27
CC METAL 77 86
CC DISULFID 77 86
CC DISULFID 99 115
CC DISULFID 114 125
CC DISULFID 148 193
CC DISULFID 182 201
CC DISULFID 224 270
CC DISULFID 269 277
CC DISULFID 289 303
CC DISULFID 302 313
CC DISULFID 340 385
CC DISULFID 384 393
CC DISULFID 416 462
CC DISULFID 461 472
CC DISULFID 485 501
CC DISULFID 500 511
CC DISULFID 538 583
CC DISULFID 582 591
CC SEQUENCE 608 AA; 68914 MW; CF5B92647AAFE9A2 CRC64;
Query Match 83.0%; Score 39; DB 1; Length 608;
Best Local Similarity 87.5%; Pred. No. 5.8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 KOEPERNE 9
Db 117 KKEPERNE 124
RESULT 13
A37C_DROSI STANDARD; PRT; 501 AA.
ID A37C_DROSI
AC O96566;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Anon-37Cs protein (Fragment).
GN CG10561 OR ANON-37CS OR CS.
OS Drosophila simulans (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=St. Lucia;
 RX MEDLINE=9250256; PubMed=10231575;
 RA Tataronkov A., Saez A.G., Ayala F.J.,
 RT "A compact gene cluster in Drosophila: the unrelated Cs gene is
 RT compressed between duplicated and Ddc.";
 RL Gene 231:111-120(1999).
 CC -1- FUNCTION: Has a nonvital function (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF091327; AAC67579.1;
 DR FlyBase; FBgn0025651; Dsim\CG10561.
 DR InterPro; IPR002937; Amino oxidase.
 DR InterPro; IPR00205; NAD_ox.
 DR Pfam; PF01593; Amino_oxidase; 1.
 FT NON_TER
 FT 1
 SQ SEQUENCE 501 AA; 56018 MW; 1F140F1E45FC1249 CRC64;
 Query Match 74.5%; Score 35; DB 1; Length 501;
 Best Local Similarity 75.0%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKQEPERN 8
 Db 352 AEQQPERN 359
 RESULT 14
 ID ALBU CHICK STANDARD; PRT; 615 AA.
 AC P19121;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).
 GN ALB.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Cassidy A.I., Salkild C.K., Baverstock P., Wallace J.C.;
 RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=83161037; PubMed=6187737;
 RA Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.;
 RT "The 5' noncoding and flanking regions of the avian very low density
 RT apolipoprotein II and serum albumin genes. Homologies with the egg
 RT white protein genes";
 RL J. Biol. Chem. 258:4556-4564(1983).
 RN [3]
 RP SEQUENCE OF 19-30.
 RX MEDLINE=78015943; PubMed=911327;
 RA Rosen A.M., Geller D.M.;
 RT "Chicken microsomal albumin: amino terminal sequence of chicken
 RT proalbumin.";

RL Biochem. Biophys. Res. Commun. 78:1060-1066(1977).
 RN [4]
 RP ALLERGENIC PROPERTIES.
 RX MEDLINE=21381307; PubMed=11488669;
 RA Quirce S., Matanon F., Umpierrez A., de las Heras M.,
 RA Fernandez-Caldas E., Sastre J.;
 RT "Chicken serum albumin (Gal d 5*) is a partially heat-labile inhalant
 RT and food allergen implicated in the bird-egg syndrome.";
 RL Allergy 56:754-762(2001).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- ALLERGEN: Causes an allergic reaction in human. Binds IgE.
 CC Partially heat-labile allergen that may cause both respiratory and
 CC food-allergy symptoms in patients with the bird-egg syndrome.
 CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -1- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X60688; CAA43098.1;
 DR EMBL; V00381; CAA23680.1;
 DR PIR; S15571; ABCHS.
 DR HSP; P02758; 1E7B.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
 FT SIGNAL 1 18
 FT PROPEP 19 23
 FT CHAIN 24 615 SERUM ALBUMIN.
 FT DOMAIN 24 209 ALBUMIN 1.
 FT DOMAIN 216 401 ALBUMIN 2.
 FT DOMAIN 408 559 ALBUMIN 3.
 FT METAL 30 30 COPPER (BY SIMILARITY).
 FT DISULFID 80 89 BY SIMILARITY.
 FT DISULFID 102 118 BY SIMILARITY.
 FT DISULFID 117 128 BY SIMILARITY.
 FT DISULFID 152 197 BY SIMILARITY.
 FT DISULFID 196 205 BY SIMILARITY.
 FT DISULFID 228 274 BY SIMILARITY.
 FT DISULFID 273 281 BY SIMILARITY.
 FT DISULFID 293 307 BY SIMILARITY.
 FT DISULFID 306 317 BY SIMILARITY.
 FT DISULFID 344 359 BY SIMILARITY.
 FT DISULFID 368 397 BY SIMILARITY.
 FT DISULFID 420 466 BY SIMILARITY.
 FT DISULFID 465 476 BY SIMILARITY.
 FT DISULFID 489 505 BY SIMILARITY.
 FT DISULFID 504 515 BY SIMILARITY.
 FT DISULFID 542 587 BY SIMILARITY.
 FT DISULFID 586 595 BY SIMILARITY.
 FT CARBOHYD 500 500 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 24 24 F -> M (IN REF. 3).
 SQ SEQUENCE 615 AA; 69918 MW; E5924BBAC066C6 CRC64;
 Query Match 74.5%; Score 35; DB 1; Length 615;
 Best Local Similarity 66.7%; Pred. No. 34;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AKQEPERNE 9

Db 119 SKADPERNE 127

RESULT 15

SUR6 XENLA STANDARD; PRT; 341 AA.
 ID -SUR6 XENLA STANDARD; PRT; 341 AA.
 AC Q9180;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 18-OCT-2001 (Rel. 40, Last sequence update)
 DT 18-OCT-2001 (Rel. 40, Last annotation update)
 DE Surfeit locus protein 6 homolog.
 GN SURF6.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wolff C.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Involved in a nucleolar function (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
 CC -!- SIMILARITY: Belongs to the SURF6 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ276843; CAB98154.1; -.
 DR InterPro; IPR007019; SURF6.
 DR Pfam; PF04935; SURF6; 1.
 KW Nuclear protein.

SQ SEQUENCE 341 AA; 39901 MW; 280A06920D653A0 CRC64;

 Query Match 72.3%; Score 34; DB 1; Length 341;
 Best Local Similarity 77.8%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
 DB 152 AAQEPPEGNE 160

Search completed: April 19, 2004, 11:52:49
 Job time : 1.65651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 3.47368 seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_92_100

Perfect score: 47

Sequence: 1 AKQEPERNE 9

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_todent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	100.0	396	Q81UK7	Q81UK7 homo sapien
2	43	91.5	205	11 Q8CG74	Q8CG74 mus musculus
3	43	91.5	576	11 Q8C7C7	Q8C7C7 mus musculus
4	43	91.5	608	5 Q95VB7	Q95VB7 schistosoma
5	43	91.5	608	11 Q8C7H3	Q8C7H3 mus musculus
6	39	83.0	584	6 Q7YSG3	Q7YSG3 felis silve
7	36	76.6	115	5 Q8MRZ4	Q8MRZ4 drosophila
8	36	76.6	203	5 Q8IN20	Q8IN20 drosophila
9	36	76.6	439	5 Q917X6	Q917X6 drosophila
10	36	76.6	456	5 Q46078	Q46078 drosophila
11	35	74.5	414	2 Q7WSR5	Q7WSR5 arthrobacte
12	35	74.5	443	11 Q8BI09	Q8BI09 mus musculus
13	35	74.5	566	11 Q91WGS	Q91WGS mus musculus
14	35	74.5	614	13 Q91134	Q91134 naja naja
15	35	74.5	884	10 Q9LQ59	Q9LQ59 arabidopsis
16	35	74.5	1853	2 Q33761	Q33761 streptococc

17	35	74.5	1854	2	Q33764	Q33764 streptococc
18	35	74.5	1854	2	Q33759	Q33759 streptococc
19	35	74.5	1874	2	Q33765	Q33765 streptococc
20	35	74.5	1874	2	Q33763	Q33763 streptococc
21	35	74.5	1878	2	Q59986	Q59986 streptococc
22	35	74.5	1914	2	Q33762	Q33762 streptococc
23	34	72.3	318	4	Q9P1N1	Q9P1N1 homo sapien
24	34	72.3	375	2	Q62290	Q62290 agrobacteri
25	34	72.3	421	11	Q88993	Q88993 mus musculu
26	34	72.3	448	11	Q99K39	Q99K39 mus musculu
27	34	72.3	684	4	Q9Y6X4	Q9Y6X4 homo sapien
28	34	72.3	750	11	Q7TM96	Q7TM96 rattus norv
29	34	72.3	772	11	Q99MD9	Q99MD9 mus musculu
30	34	72.3	773	11	Q35499	Q35499 mus musculu
31	34	72.3	2011	4	Q920A9	Q920A9 homo sapien
32	33	70.2	65	4	Q9BUN3	Q9BUN3 homo sapien
33	33	70.2	77	15	Q7ZPH8	Q7ZPH8 human immun
34	33	70.2	178	12	Q7TD10	Q7TD10 cestrum yel
35	33	70.2	245	16	Q8X621	Q8X621 escherichia
36	33	70.2	256	16	Q8FB79	Q8FB79 escherichia
37	33	70.2	284	16	Q7VUL4	Q7VUL4 bordetella
38	33	70.2	343	10	Q9S1Z8	Q9S1Z8 arabidopsis
39	33	70.2	399	16	Q8K6Q6	Q8K6Q6 bacillus ha
40	33	70.2	407	16	Q91323	Q91323 pseudomonas
41	33	70.2	433	16	Q9HTB4	Q9HTB4 pseudomonas
42	33	70.2	503	5	Q25777	Q25777 plasmodium
43	33	70.2	521	5	Q9VDD3	Q9VDD3 drosophila
44	33	70.2	555	4	Q8NA99	Q8NA99 homo sapien
45	33	70.2	713	16	Q8UUI6	Q8UUI6 agrobacteri

ALIGNMENTS

RESULT 1
Q81UK7 PRELIMINARY; PRT; 396 AA.

AC Q81UK7;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1; -
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 47; DB 4; Length 396;
Best Local Similarity 100.0%; Pred No. 0.43; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

Qy 1 AKQEPERNE 9
Db 116 AKQEPERNE 124

RESULT 2
Q8CG74

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KQEPERNE 9
 Db 85 KQEPERNE 92

RESULT 4

Q95VB7
 ID Q95VB7 PRELIMINARY; PRT; 608 AA.
 AC Q95VB7;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Albumin.
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
 OC Schistosomatoidea; Schistosomatidae; Schistosoma.
 OX NCBI_TaxID=6183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
 RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
 component in murine infection with Schistosoma mansoni."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF418550; AAL08579.1; -
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; P:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 2.
 DR SEQUENCE 608 AA; 68225 MW; ESEABB28E1C66E54 CRC64;

Query Match 91.5%; Score 43; DB 5; Length 608;
 Best Local Similarity 88.9%; Pred. No. 4.2;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
 Db 116 AKQEPERNE 124

RESULT 5

Q8C7H3
 ID Q8C7H3 PRELIMINARY; PRT; 608 AA.
 AC Q8C7H3;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Albumin 1.
 GN ALB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Liver;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 the RIKEN Genome Exploration Research Group Phase I & II Team;
 Nature 420:563-573 (2002)."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK050248; BAC34145.1; -
 DR MGD; MGI:87991; Alb1.
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; P:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.

Query Match 91.5%; Score 43; DB 11; Length 576;
 Best Local Similarity 100.0%; Pred. No. 3.9;

QY 2 KQEPERNE 9
 Db 117 KQEPERNE 124

RESULT 3

Q8C7C7
 ID Q8C7C7 PRELIMINARY; PRT; 576 AA.
 AC Q8C7C7;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Albumin 1 (Fragment).
 GN ALB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 the RIKEN Genome Exploration Research Group Phase I & II Team;
 Nature 420:563-573 (2002)."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK050644; BAC34360.1; -
 DR MGD; MGI:87991; Alb1.
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; P:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 DR NON TER 1
 FT SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;

Query Match 91.5%; Score 43; DB 11; Length 576;
 Best Local Similarity 100.0%; Pred. No. 3.9;

DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transprot_prot; 3.
DR PRINTS: P00802; SERUMALBUMIN.
DR PRODOM: PD002486; Serum_albumin; 1.
DR SMART: SM00103; ALBUMIN; 3.
DR PROSITE: PS0012; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600ED3A61B4 CRC64;
Query Match 91.5%; Score 43; DB 11; Length 608;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 KOEPERNE 9
Db 117 KOEPERNE 124
RESULT 6
QYSG3 QYSG3 PRELIMINARY; PRT; 584 AA.
AC QYSG3
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Albumin (fragment).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
albumin: IgE recognition, induction of basophil activation and
lymphoproliferative responses in atopic patients.";
RL Submitted (May 2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ487677; CAD32275.1; -.
FT NON TER 1
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;
Query Match 83.0%; Score 39; DB 6; Length 584;
Best Local Similarity 87.5%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 KOEPERNE 9
Db 93 KOEPERNE 100
RESULT 7
QYSG3 QYSG3 PRELIMINARY; PRT; 115 AA.
AC QYSG3
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)
DE SD11171p.
GN BCDNA:SD11171.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuno J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Celniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY119187; AAMS1047.1; -.
DR FlyBase: FBGN0062887; BCDNA:SD11171.
SQ SEQUENCE 115 AA; 13534 MW; F680D64C7C967AEC CRC64;
Query Match 76.6%; Score 36; DB 5; Length 115;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 KOEPERNE 9
Db 34 KOEPERNE 41
RESULT 8
QYSG3 QYSG3 PRELIMINARY; PRT; 203 AA.
AC QYSG3
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE CG31174-PA.
GN CG31174.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Frankoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodgett W., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.H.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,

RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Matzel B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,
RT "Sequencing of Drosophila melanogaster genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.W.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003736; AAN13887.1; -;
DR FlyBase; FBgn0051174; CG31174.
SQ SEQUENCE 203 AA; 23859 MW; 3C3BEF93F3434B93 CRC64;

Query Match 76.6%; Score 36; DB 5; Length 203;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9
:|||||:
DB 34 KREPERNQ 41

RESULT 9
Q917X6 PRELIMINARY; PRT; 439 AA.
ID Q917X6
AC Q917X6
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE EG:39E1.1 protein (LD42227p).
GN EG:39E1.1 OR CG11596.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Baldwin D.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleissmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Harvey D., Helman I.J., Hernandez J.R., Houck J.,
RA Hostin D., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kessel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzay D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003422; AAG22387.1; -;
DR EMBL: AY051951; AAK93375.1; -;
DR FlyBase; FBgn0023522; EG:39E1.1.
SQ SEQUENCE 439 AA; 50504 MW; 883DE54CC8D8B5F7 CRC64;

Query Match 76.6%; Score 36; DB 5; Length 439;
Best Local Similarity 75.0%; Pred. No. 74;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPERNE 9
:|||||:
DB 408 EQPERNE 415

RESULT 10
O46078 PRELIMINARY; PRT; 456 AA.
ID O46078
AC O46078; Q9M540;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE EG:39E1.1 protein.
GN EG:39E1.1 OR CG11596.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Baldwin D.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

425 EQPERNE 432

Db

RESULT 11

Q7WSR5 PRELIMINARY; PRT; 414 AA.

ID Q7WSR5; 01-OCT-2003 (TrEMBLrel. 25, Created)

AC Q7WSR5; 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative transmembrane protein.

OS Arthrobracter ilicis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Micrococcinea; Micrococcaceae; Arthrobracter.

NCBI_taxid=43665;

NCBI_taxid=43665;

SEQUENCE FROM N.A.

STRAIN=Rue61a;

RA MEDLINE=22753791; Pubmed=12730200;

RT "Gagne K., Hauer B., Kraft R., Ruettermann J., Fetzner S.;

RT of Quinaldine to Anthranilate, Characterization and Functional

RT Expression of the Quinaldine 4-oxidase qoxLMS Genes.";

RL J. Biol. Chem. 278:27483-27494 (2003).

DR EMBL; AJ537472; CAD61035.1; .

KW Monooxygenase; Transmembrane.

SQ SEQUENCE 414 AA; 45307 MW; 00BD746F673C0486 CRC64;

Query Match 74.5%; Score 35; DB 2; Length 414;

Best Local Similarity 66.7%; Pred. No. 1.1e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKOEPPERNE 9

DB 146 ARQDPERRE 154

RESULT 12

Q8BIQ9 PRELIMINARY; PRT; 443 AA.

AC Q8BIQ9; 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DE Similar to AMP activated protein kinase gamma 1.

GN 2410051C13RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

NCBI_taxid=10090;

NCBI_taxid=10090;

SEQUENCE FROM N.A.

STRAIN=C57BL/6J; TISSUE=Olfactory brain;

RA MEDLINE=22354683; Pubmed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).

DR EMBL; AK032238; BAC27775.1; .

DR MGD; MGI:1920950; 2410051C13RIK.

DR InterPro; IPR000644; CBS_domain.

DR Pfam; PF00571; CBS; 4.

DR SMART; SM00116; CBS; 4.

SQ SEQUENCE 443 AA; 49980 MW; AC336EAD3FE35E28 CRC64;

Query Match 74.5%; Score 35; DB 11; Length 443;

Best Local Similarity 87.5%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KOEPPERNE 9

DB 57 KOEPPERNE 64

Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

Borkova D., Botchan M.R., Bouck J., Brckstein P., Brothier P.,

Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

Kimmel B.E., Kodira C.D., Kraft A., Kravitz S., Kulp D., Lai Z.,

Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,

Liu X., Mattet B., McIntosh T.C., McLeod M.P., McPherson D.,

Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacieb J.M.,

Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,

Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,

Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

"The genome sequence of Drosophila melanogaster.";

Science 287:2185-2195 (2000).

SEQUENCE FROM N.A.

STRAIN=OREGON-R;

RA MEDLINE=20196011; Pubmed=10731137;

RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D., Batrell B.,

Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E., Dreano S.,

Gloux S., Lelaure V., Mottier S., Gallibert F., Borkova D., Minana B.,

Kafatos F.C., Louis C., Siden-Klamos I., Bolshakov S.,

Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,

Modellell J., Peter A., Schottler P., Werner M., Mourikioti P.,

Beinert N., Dowe G., Schafer U., Jackle H., Bucheton A.,

Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,

McMillan P.J., Salles C., Tait E.A., Valenti P., Saunderson R.D.,

Glover D.M.;

"From sequence to chromosome: the tip of the X chromosome of D.

melanogaster.";

Science 287:2220-2222 (2000).

SEQUENCE 456 AA; 52441 MW; CC6BCP2FBADD5954 CRC64;

Name=long;

Isoid=046078-1; Sequence=Displayed;

Name=short;

Isoid=046078-2; Sequence=VSP_050174;

EMBL; AE003422; AAF45697.1; .

EMBL; AE003422; AAF45698.1; .

EMBL; AL009191; CAA15684.1; .

DR Flybase; FBgn0023522; EG:39E1.1.

Alternative splicing; Hypothetical protein.

Missing (in isoform short).

FTID=VSP_050174.

VARSPPLIC 1 55

SEQUENCE 456 AA; 52441 MW; CC6BCP2FBADD5954 CRC64;

Query Match 76.6%; Score 35; DB 5; Length 456;

Best Local Similarity 75.0%; Pred. No. 77;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KOEPPERNE 9

DB 11 KOEPPERNE 64


```
RESULT 13
Q91WGS
ID Q91WGS PRELIMINARY; PRT; 566 AA.
AC Q91WGS;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN 2410051C13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015283; AAH15283.1; ...
DR MGD; MGI:1920950; 2410051C13RIK.
DR InterPro; IPR000644; CBS_domain.
DR Pfam; PF00571; CBS; 4.
DR SMART; SM00116; CBS; 4.
KW Hypothetical protein.
SQ SEQUENCE 566 AA; 62997 MW; 754DE4D696102C7D CRC64;

Query Match 74.5%; Score 35; DB 11; Length 566;
Best Local Similarity 87.5%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KOEPERNE 9
DB 181 KOEPERPE 188

RESULT 14
Q91134
ID Q91134 PRELIMINARY; PRT; 614 AA.
AC Q91134;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cobra serum albumin.
OS Naja naja (Indian cobra).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
OX NCBI_TaxID=35670;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96145734; PubMed=8561913;
RA Wang X., Hansen H., Havsteen B.;
RT "Evidence of the coevolution of snake toxin and its enogenous
RT antioxin. Cloning, sequence and expression of a serum albumin cDNA of
RT the chinese cobra.";
RL Biol. Chem. Hoppe-Seyler 376:545-553 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93343093; PubMed=8343135;
RA Shao J., Shen H., Havsteen B.;
RT "Purification, characterization and binding interactions of the
RT Chinese-cobra (Naja naja atra) serum antitoxic protein CSAP.";
RL Biochem. J. 293:559-566 (1993).
DR EMBL; X78598; CAA55333.1; ...
DR HSSP; P02768; 1E7B.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0004810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
```

```
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 614 AA; 69798 MW; 3DB2D3CC4BD8CBFD CRC64;

Query Match 74.5%; Score 35; DB 13; Length 614;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
DB 119 AKADPERNE 127

RESULT 15
Q9LQ99
ID Q9LQ99 PRELIMINARY; PRT; 884 AA.
AC Q9LQ99;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE T4012.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Chao Q., Shinn P., Brooks S., Buehler E., Dunn P., Khan S., Kim C.,
RA Walker M., Brooks S., Altafi H., Araujo R., Conn L., Conway A.B.,
RA Gonzalez A., Hansen N.F., Ruizar L., Kremenetskaia I., Lenz C., Li J.,
RA Liu S., Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V.,
RA Yu G., Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC T4012 from chromosome
RT I.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chlou J., Choi E., Conn L.,
RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chlou J., Choi E.,
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007396; AAF26771.2; ...
DR PIR; G96787; G96787.
DR InterPro; IPR005378; Vps35.
DR Pfam; PF03635; Vps35; 1.
SQ SEQUENCE 884 AA; 100411 MW; B4BFD781BDALD048 CRC64;
```

Query Match 74.5%; Score 35; DB 10; Length 884;
Best Local Similarity 87.5%; Fred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QV 2 KQEPERNE 9
|||
Db 233 KQEKERNE 240
|||

Search completed: April 19, 2004, 12:00:03
Job time : 5.47368 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:23 ; Search time 5.29363 Seconds
(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_92_100

Perfect score: 47
Sequence: 1 AKQEPERNE 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	116	AAU29580	Novel hum
2	47	100.0	116	AAU29574	Novel hum
3	47	100.0	137	AAO12088	Human pol
4	47	100.0	143	ABP05068	Human ORF
5	47	100.0	156	AAU33073	Novel hum
6	47	100.0	156	AAU33271	Novel hum
7	47	100.0	192	AAU29875	Novel hum
8	47	100.0	195	AAO17048	Human ser
9	47	100.0	195	ABU10022	Human ser
10	47	100.0	204	AAU83947	Yeast cod
11	47	100.0	236	AAO17051	Human alb
12	47	100.0	241	AAO16984	Alpha-MSH
13	47	100.0	241	ABU10025	Alpha-MSH
14	47	100.0	242	AAO16985	Alpha-MSH
15	47	100.0	244	AAO16986	Alpha-MSH
16	47	100.0	245	AAO16987	Alpha-MSH
17	47	100.0	245	AAO16988	Alpha-MSH
18	47	100.0	268	AAO16989	Alpha-MSH
19	47	100.0	268	ABU10026	Alpha-MSH
20	47	100.0	303	AAU14178	Human ser
21	47	100.0	327	AAU32564	Novel hum
22	47	100.0	327	AAU29942	Novel hum
23	47	100.0	327	AAU32995	Novel hum
24	47	100.0	327	AAU33287	Novel hum
25	47	100.0	373	AAU90387	N-termina

26	47	100.0	388	1	AAU90389	Novel hum
27	47	100.0	389	1	AAU90390	N-termina
28	47	100.0	390	1	AAU90391	N-termina
29	47	100.0	407	1	AAU90392	N-termina
30	47	100.0	500	7	ADD32019	Heterolog
31	47	100.0	584	6	ABG72381	Mature hu
32	47	100.0	585	1	AAU90344	Sequence
33	47	100.0	585	1	AAU90388	Mature hu
34	47	100.0	585	1	AAU91422	Human nor
35	47	100.0	585	1	AAU90318	Human ser
36	47	100.0	585	2	AAU08457	Human ser
37	47	100.0	585	2	AAU26207	Human ser
38	47	100.0	585	2	AAU26362	Synthetic
39	47	100.0	585	2	AAU20029	Human ser
40	47	100.0	585	2	AAU80301	Human ser
41	47	100.0	585	2	AAU20111	HSA prote
42	47	100.0	585	2	AAU59841	Mature pr
43	47	100.0	585	3	AAU84873	Amino aci
44	47	100.0	585	3	AAU83946	Yeast cod
45	47	100.0	585	4	AAU52567	Mature hu

ALIGNMENTS

RESULT 1
AAU29580
ID AAU29580 standard; protein; 116 AA.

XX AAU29580;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #71.

XX Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

PN 25-OCT-2001.

XX 16-APR-2001; 2001WO-US008656.

XX 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic

XX vaccination, testing and therapy.

XX Claim 20; Page 169; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell

CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 116 AA;

Query Match 100.0%; Score 47; DB 4; Length 116;
 Best Local Similarity 100.0%; Pred. No. 0.49; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

Qy 1 AKQEPERNE 9
 |||||
 Db 6 AKQEPERNE 14

RESULT 2

AAU29574
 ID AAU29574 standard; protein; 116 AA.

XX AC AAU29574;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human secreted protein #55.

XX KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX XX WO200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US008656.

XX PR 18-APR-2000; 2000US-00552929.

XX PR 26-JAN-2001; 2001US-00770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX XX WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.

XX PS Claim 20; Page 167; 765pp; English.

XX CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX

XX SQ Sequence 116 AA;

Query Match 100.0%; Score 47; DB 4; Length 116;
 Best Local Similarity 100.0%; Pred. No. 0.49; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

Qy 1 AKQEPERNE 9
 |||||
 Db 6 AKQEPERNE 14

RESULT 3

AAO12088
 ID AAO12088 standard; protein; 137 AA.

XX AC AAO12088;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 25980.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX XX (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AAI92019.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 20; SEQ ID NO 25980; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 137 AA;

Query Match 100.0%; Score 47; DB 4; Length 137;
 Best Local Similarity 100.0%; Pred. No. 0.58;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
 |||||
 Db 121 AKQEPERNE 129

RESULT 4

ABP05068	ABP05068 standard; protein; 143 AA.
XX	
AC	ABP05068;
XX	
DT	24-JUN-2002 (first entry)
XX	
DE	Human ORFX protein sequence SEQ ID NO:10118.
XX	
KW	Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW	hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW	degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW	cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW	hypertension; hypothyroidism; cholesterol ester storage disease;
KW	immune deficiency; immune disorder; infectious disease;
KW	autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW	myasthenia gravis.
XX	
OS	Homo sapiens.
XX	
PN	WO200192523-A2.
FD	06-DEC-2001.
XX	
PF	29-MAY-2001; 2001WO-US010836.
XX	
PR	30-MAY-2000; 2000US-0206132P.
DR	29-AUG-2000; 2000US-0228716P.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Shimkets RA, Leach MD;
XX	
DR	WPI; 2002-106308/14.
DR	N-PSDB; ABN20820.
XX	
PT	Novel human polypeptides and polynucleotides useful for diagnosing,
PT	preventing and treating cardiovascular disease, neurodegenerative,
PT	hyperproliferative disorders and autoimmune disorders.
XX	
PS	Disclosure; SEQ ID NO 10118; 1037pp; English.
XX	
CC	The present invention describes substantially purified human proteins
CC	(referred to as open reading frame, ORFX, where X is 1-11491 (see Table
CC	in the specification)). AEN15762 to AEN27952 encode the human ORFX
CC	proteins given in ABP00010 to ABP1500. ORFX proteins are useful for
CC	treating or preventing a pathology associated with an ORFX-associated
CC	disorder in humans, and in the manufacture of a medicament for treating
CC	syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC	sequences can be used in gene therapy. ORFX sequences can be used in the
CC	treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC	psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC	osteoarthritis, neurodegenerative disorders, disorders related to organ
CC	transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC	lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC	storage disease, various immune deficiencies and disorders, infectious
CC	diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC	arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC	disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC	useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC	bone degenerative disorders, or periodontal disease, and for gut
CC	protection or regeneration and treatment of lung or liver fibrosis,
CC	repfusion injury in various tissues and conditions resulting from
CC	systemic cytokine damage. N.B. the sequence data for this patent did not
CC	form part of the printed specification, but was obtained in electronic
CC	format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 143 AA;

Qy	1	AKQEPERNE 9	
Db	111	AKQEPERNE 119	
RESULT 5			
AAU33073			
ID	AAU33073	standard; protein; 156 AA.	
XX	AAU33073;		
XX	18-DEC-2001	(first entry)	
XX	Novel human secreted protein #3564.		
XX	Human; vaccination; gene therapy; nutritional supplement;		
KW	stem cell proliferation; haematopoiesis; nerve tissue regeneration;		
KW	immune suppression; immune stimulation; anti-inflammatory; leukaemia.		
XX			
OS	Homo sapiens.		
XX			
XX	WO200179449-A2.		
XX	25-OCT-2001.		
XX	16-APR-2001; 2001WO-US008656.		
XX	18-APR-2000; 2000US-Q0552929.		
XX	26-JAN-2001; 2001US-00770160.		
XX	(HYSE-) HYSEQ INC.		
PA			
XX			
PI	Tang YT, Liu C, Drmanac RT;		
XX	WPI; 2001-611725/70.		
XX			
PT	Nucleic acids encoding a range of human polypeptides, useful in genetic		
XX	vaccination, testing and therapy.		
XX			
PS	Claim 20; Page 704; 765pp; English.		
XX			
CC	The invention relates to novel human secreted polypeptides. The		
CC	polypeptides and antibodies to the polypeptides are useful for		
CC	determining the presence of or predisposition to a disease associated		
CC	with altered levels of polypeptide. The polypeptides are also useful for		
CC	identifying agents (agonists and antagonists) that bind to them. Cells		
CC	expressing the proteins are useful for identifying a therapeutic agent		
CC	for use in treatment of a pathology related to aberrant expression or		
CC	physiological interactions of the polypeptide. Vectors comprising the		
CC	nucleic acids encoding the polypeptides and cells genetically engineered		
CC	to express them are also useful for producing the proteins. The proteins		
CC	are useful in genetic vaccination, testing and therapy, and can be used		
CC	as nutritional supplements. They may be used to increase stem cell		
CC	proliferation; to regulate haematopoiesis; and in bone, cartilage, tendo		
CC	and/or nerve tissue growth or regeneration; immune suppression and/or		
CC	stimulation, as anti-inflammatory agents; and in treatment of leukaemias		
CC	AAU29510-AAU33304 represent the amino acid sequences of novel human		
CC	secreted proteins of the invention		
XX			
SQ	Sequence 156 AA;		

```

Query Match      100.0%; Score 47; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 0.66;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AKQEPNER 9
         |||||
Db      47 AKQEPNER 55

```

```
Query Match      100.0%; Score 47; DB 5; Length 143;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Query Match 100.0%; Score 4/; DB 5; Length 143;
Best Local Similarity 100.0%; Pred. No. 0.61;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
AAU33271
ID AAU33271 standard; protein: 156 AA.

AAU33271
ID AAU33271 standard; protein; 156 AA.

XX AC AAU3327L;
 XX DT 18-DEC-2001 (first entry)
 XX DE
 XX DE Novel human secreted protein #3762.
 XX KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX OS Homo sapiens.
 XX PN WO200179449-A2.
 XX PD 25-OCT-2001.
 XX PF 16-APR-2001; 2001WO-US008656.
 XX PR 18-APR-2000; 2000US-00552929.
 XX PR 26-JAN-2001; 2001US-00770160.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX DR WPI; 2001-611725/70.
 XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX PS Claim 20; Page 751; 765pp; English.
 XX CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX SQ Sequence 156 AA;
 Query Match 100.0%; Score 47; DB 4; Length 156;
 Best Local Similarity 100.0%; Pred. No. 0.66;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKQEPERNE 9
 |||||
 Db 47 AKQEPERNE 55
 RESULT 7
 AAU29875
 ID AAU29875 standard; protein; 192 AA.
 XX AC AAU29875;
 XX DE
 XX DE 18-DEC-2001 (first entry)
 XX DE Novel human secreted protein #366.
 XX KW Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 OS Homo sapiens.
 PN WO200179449-A2.
 PD 25-OCT-2001.
 PF 16-APR-2001; 2001WO-US008656.
 XX PR 18-APR-2000; 2000US-00552929.
 XX PR 26-JAN-2001; 2001US-00770160.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX DR WPI; 2001-611725/70.
 XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX PS Claim 20; Page 206; 765pp; English.
 XX CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX SQ Sequence 192 AA;
 Query Match 100.0%; Score 47; DB 4; Length 192;
 Best Local Similarity 100.0%; Pred. No. 0.81;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKQEPERNE 9
 |||||
 Db 98 AKQEPERNE 106
 RESULT 8
 AAU17048
 ID AAU17048 standard; protein; 195 AA.
 XX AC AAU17048;
 XX DE
 XX DE 29-MAY-2002 (first entry)
 XX DE Human serum albumin (1-195) SEQ ID NO: 56.
 XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
 KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritis;
 KW antidiabetic; antibacterial; dermatological; antiporiatic;
 KW diabetes; uveitis; coeliac disease.
 OS Homo sapiens.

XX WO200206316-A2.
 XX 24-JAN-2002.
 XX 16-JUL-2001; 2001WO-US022263.
 XX 14-JUL-2000; 2000US-0218381P.
 XX 18-AUG-2000; 2000US-0226382P.
 XX 06-OCT-2000; 2000US-0238380P.
 XX 29-DEC-2000; 2000US-0258764P.
 XX 14-JUN-2001; 2001US-0298317P.
 XX (ZYCO-) ZYCO INC.
 XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
 XX WPI; 2002-195901/25.
 XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
 XX stimulating hormone concatamer or its analog, for treating inflammatory
 XX or autoimmune disorders.
 XX Example 2; Page 46; 89pp; English.
 XX The present invention relates to a nucleic acid comprising a sequence
 XX encoding a fusion polypeptide having an alpha-melanocyte stimulating
 XX hormone (MSH) concatamer. The sequences are useful for treating an
 XX individual suffering from, or at risk of, a disorder of the immune system
 XX e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
 XX arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
 XX hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
 XX multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
 XX present sequence is a protein described in the exemplification of the
 XX invention
 XX SQ Sequence 195 AA;
 Query Match 100.0%; Score 47; DB 5; Length 195;
 Best Local Similarity 100.0%; Pred. No. 0.82;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKQEPERNE 9
 DB 92 AKQEPERNE 100
 RESULT 9
 ABU10022
 ID ABU10022 standard; protein; 195 AA.
 XX AC ABU10022;
 XX 31-JUL-2003 (first entry)
 XX Human serum albumin residues 1-195.
 XX Bladder disorder; cytostatic; antiinflammatory; immune response;
 XX un-methylated CpG sequence; alpha-MSH; melanocortin receptor;
 XX bladder cancer; tumour; interstitial cystitis; inflammation;
 XX alpha-MSH concatamer; melanocyte stimulating hormone; human;
 XX serum albumin.
 XX Homo sapiens.
 XX US2002193332-A1.
 XX 19-DEC-2002.
 XX 12-FEB-2002; 2002US-00074956.
 XX 12-FEB-2001; 2001US-0268175P.
 XX

PA (HEDL/) HEDLEY M L.
 XX Hedley ML;
 XX WPI; 2003-447327/42.
 XX Modulating immune responses in a mammal with a bladder disorder e.g.
 XX bladder cancer, by administering nucleic acids comprising un-methylated
 XX CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to
 XX the mammal.
 XX Example 2; Page 9; 17pp; English.
 XX The invention describes a method of modulating an immune response in a
 XX mammal, comprising identifying a mammal that has or is at risk for having
 XX a bladder disorder, and administering: (a) an isolated nucleic acid (N1)
 XX comprising an un-methylated CpG sequence to the mammal, (b) an isolated
 XX nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal,
 XX or (c) a peptide that binds to a melanocortin receptor to the mammal. The
 XX method is useful for modulating immune response in a mammal having a
 XX bladder disorder, where administration of (N1) results in an amelioration
 XX of one or more symptoms of the disorder. Preferably, the method is useful
 XX for modulating immune response in a mammal having bladder cancer (where
 XX administration of (N1) results in a decrease in tumour size or activity),
 XX or for modulating immune response in a mammal having interstitial
 XX cystitis (where administration of (N1) results in a modulation of the
 XX immune response from Th2 response to a Th1 response). The method is also
 XX useful for modulating immune response in a mammal having bladder disorder
 XX that is characterised by inflammation which is associated with symptoms
 XX of interstitial cystitis or associated with a disruption of the integrity
 XX of the bladder lining. This is the amino acid sequence of human serum
 XX albumin residues 1-195 that can be used in the creation of melanocyte
 XX stimulating hormone (alpha-MSH) concatamers resulting in secretion of the
 XX fusion protein when expressed in mammalian cells
 XX SQ Sequence 195 AA;
 Query Match 100.0%; Score 47; DB 7; Length 195;
 Best Local Similarity 100.0%; Pred. No. 0.82;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKQEPERNE 9
 DB 92 AKQEPERNE 100
 RESULT 10
 AAY83947
 ID AAY83947 standard; protein; 204 AA.
 XX AC AAY83947;
 XX 28-JUL-2000 (first entry)
 XX Yeast codon-biased recombinant HSA protein fragment HSA-I.
 XX Recombinant; human serum albumin; HSA; Yeast codon bias; host cell;
 XX overlapping oligonucleotide; expression vector.
 XX Homo sapiens.
 XX OS Synthetic.
 XX CN1239103-A.
 XX 22-DEC-1999.
 XX 17-JUN-1998; 98CN-00102506.
 XX 17-JUN-1998; 98CN-00102506.
 XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.
 XX Li S, Lu D;
 XX

XX WPI; 2000-351198/31.
DR N-PSDB; AAA10092.
XX
XX
XX Process for preparing recombinant human serum albumin comprising yeast
PT biased sex codons - uses a recombinant DNA technique.
XX
XX Example 1; Fig 3; 44pp; Chinese.
XX
XX The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
CC three synthetic fragments (AAA10092-Al0094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the sequence of the HSA fragment HSA-1 encoded by the human gene with a
CC yeast codon bias. The invention also covers a recombinant expression
CC vector, yeast host cells carrying the recombinant expression vector and
CC the process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
XX Sequence 204 AA;
SQ

Query Match 100.0%; Score 47; DB 3; Length 204;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
|||
99 AKQEPERNE 107

Db

RESULT 11
AAO17051
ID AAO17051 standard; protein; 236 AA.
AC AAO17051;
XX
XX 29-MAY-2002 (first entry)
XX
XX Human albumin-thrombin-alpha-MSH SEQ ID NO: 82.
XX
XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW chondros; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; aniarthritic;
KW antidiabetic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
XX Homo sapiens.
OS
XX
XX WO200206316-A2.
PN
XX
XX 24-JAN-2002.
PD
XX
XX 16-JUL-2001; 2001WO-US022263.
PF
XX
XX 14-JUL-2000; 2000US-0218381P.
PR
XX 18-AUG-2000; 2000US-0228382P.
PR
XX 06-OCT-2000; 2000US-0238380P.
PR
XX 29-DEC-2000; 2000US-0258764P.
PR
XX 14-JUN-2001; 2001US-0298317P.
XX
XX (ZYCO-) ZYCOS INC.
PA
XX
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
PI
XX WPI; 2002-195801/25.
DR
XX
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concanamer or its analog, for treating inflammatory
PT or autoimmune disorders.
PT

CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention

XX SQ Sequence 241 AA;
Query Match 100.0%; Score 47; DB 5; Length 241;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKOEPERNE 9
Db 116 AKOEPERNE 124
|||||

RESULT 13
ABU10025
ID ABU10025 standard; protein; 241 AA.
XX AC ABU10025;
XX 31-JUL-2003 (first entry)
XX Alpha-MSH/Serum albumin fusion protein H9.
XX Bladder disorder; cytostatic; antiinflammatory; immune response;
XX un-methylated CpG sequence; alpha-MSH; melanocortin receptor;
XX bladder cancer; tumour; interstitial cystitis; inflammation;
XX alpha-MSH concatamer; melanocyte stimulating hormone; human;
XX serum albumin; fusion protein; H9.
XX Homo sapiens.
OS Synthetic.
XX US2002193332-A1.
XX 19-DEC-2002.
XX 12-FEB-2002; 2002US-00074956.
XX 12-FEB-2001; 2001US-0268175P.
XX (HEDL/) HEDLEY M L.
XX Hedley ML;
XX WPI; 2003-447327/42.
XX Modulating immune responses in a mammal with a bladder disorder e.g.
PT bladder cancer, by administering nucleic acids comprising un-methylated
PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to
PT the mammal.
XX Example 2; Page 9; 17pp; English.

XX The invention describes a method of modulating an immune response in a
CC mammal, comprising identifying a mammal that has or is at risk for having
CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)
CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated
CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;
CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The
CC method is useful for modulating immune response in a mammal having a
CC bladder disorder, where administration of (N1) results in an amelioration
CC of one or more symptoms of the disorder. Preferably, the method is useful
CC for modulating immune response in a mammal having bladder cancer (where
CC administration of (N1) results in a decrease in tumour size or activity),
CC or for modulating immune response in a mammal having interstitial
CC cystitis (where administration of (N1) results in a modulation of the
CC immune response from Th2 response to a Th1 response). The method is also
CC useful for modulating immune response in a mammal having bladder disorder

CC that is characterised by inflammation which is associated with symptoms
CC of interstitial cystitis or associated with a disruption of the integrity
CC of the bladder lining. This is the amino acid sequence of a human serum
CC melanocyte stimulating hormone (alpha-MSH) concatamer useful in the
CC treatment of bladder disorders

XX SQ Sequence 241 AA;
Query Match 100.0%; Score 47; DB 7; Length 241;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKOEPERNE 9
Db 116 AKOEPERNE 124
|||||

RESULT 14
AAO16985
ID AAO16985 standard; protein; 242 AA.
XX AC AAO16985;
XX 29-MAY-2002 (first entry)
XX Alpha-MSH construct protein fragment SEQ ID NO: 70.
XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
XX alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
XX cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
XX immunosuppressive; antiinflammatory; antirheumatic; antiarthritis;
XX antidiabetic; antibacterial; dermatological; antipsoriatic;
XX antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
XX diabetes; uveitis; coeliac disease.
XX Unidentified.
OS WO200206316-A2.
XX 24-JAN-2002.
XX 16-JUL-2001; 2001WO-US022263.
XX 14-JUL-2000; 2000US-0218381P.
XX 18-AUG-2000; 2000US-0226382P.
XX 06-OCT-2000; 2000US-0238380P.
XX 29-DEC-2000; 2000US-0258764P.
XX 14-JUN-2001; 2001US-0298317P.
XX (ZYCO-) ZYCOS INC.
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX WPI; 2002-195801/25.
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX Disclosure; Page 5; 89pp; English.

XX The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention

XX SQ Sequence 242 AA;

Db 116 AKQEPERNE 124
Search completed: April 19, 2004, 11:51:17
Job time : 5.29363 secs

Query Match 100.0%; Score 47; DB 5; Length 242;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
Db 116 AKQEPERNE 124

RESULT 15
AAO16986
ID AAO16986 standard; protein; 244 AA.
XX AAO16986;
AC AAO16986;
XX AAO16986;
DT 29-MAY-2002 (first entry)
XX Alpha-MSH construct protein fragment SEQ ID NO: 71.
XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX Unidentified.
OS
XX WO200206316-A2.
PN
XX 24-JAN-2002.
PD
XX 16-JUL-2001; 2001WO-US022263.
PF
XX 14-JUL-2000; 2000US-0218381P.
PR
XX 18-AUG-2000; 2000US-0228382P.
PR
XX 06-OCT-2000; 2000US-0238380P.
PR
XX 29-DEC-2000; 2000US-0258764P.
PR
XX 14-JUN-2001; 2001US-0298317P.
XX (ZYCO-) ZYCOS INC.
PA
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
PI WPI; 2002-195801/25.
DR
XX

PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
stimulating hormone concatamer or its analog, for treating inflammatory
or autoimmune disorders.
PT
XX
PS Disclosure; Page 5; 89pp; English.
XX

CC The present invention relates to a nucleic acid comprising a sequence
encoding a fusion polypeptide having an alpha-melanocyte stimulating
hormone (MSH) concatamer. The sequences are useful for treating an
individual suffering from, or at risk of, a disorder of the immune system
e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
present sequence is a protein described in the exemplification of the
invention
CC
XX

SQ Sequence 244 AA;

Query Match 100.0%; Score 47; DB 5; Length 244;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 3.78947 Seconds
(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_92_100

Perfect score: 47

Sequence: 1 AKQEPERNE 9

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pcp.*
- 13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pcp.*
- 17: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pcp.*
- 18: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	195	13	US-10-074-956-24
2	47	100.0	241	13	US-10-074-956-27
3	47	100.0	268	13	US-10-074-956-28
4	47	100.0	585	9	US-09-929-552-2
5	47	100.0	585	10	US-09-932-613-445
6	47	100.0	585	10	US-09-984-010-26
7	47	100.0	585	10	US-09-833-117-18
8	47	100.0	585	10	US-09-833-117-18
9	47	100.0	585	10	US-09-932-322-445
10	47	100.0	585	10	US-09-832-501-18
11	47	100.0	585	11	US-09-833-118-18
12	47	100.0	585	11	US-09-833-245-18
13	47	100.0	585	12	US-10-424-999-11
14	47	100.0	585	12	US-10-425-000-31
15	47	100.0	585	12	US-10-433-108-34

16	47	100.0	585	13	US-10-153-064-5	Sequence 5, Appli
17	47	100.0	585	14	US-10-153-604A-5	Sequence 5, Appli
18	47	100.0	585	14	US-10-319-263-1	Sequence 2, Appli
19	47	100.0	585	14	US-10-319-263-2	Sequence 1, Appli
20	47	100.0	585	14	US-10-414-489-1	Sequence 2, Appli
21	47	100.0	585	14	US-10-414-489-2	Sequence 1, Appli
22	47	100.0	585	14	US-10-413-831-1	Sequence 2, Appli
23	47	100.0	585	14	US-10-413-831-2	Sequence 1, Appli
24	47	100.0	585	15	US-10-413-832-1	Sequence 2, Appli
25	47	100.0	585	15	US-10-413-832-2	Sequence 1, Appli
26	47	100.0	585	15	US-10-414-386-1	Sequence 2, Appli
27	47	100.0	585	15	US-10-414-386-2	Sequence 1, Appli
28	47	100.0	585	15	US-10-233-675A-11	Sequence 11, Appli
29	47	100.0	585	15	US-10-462-262-26	Sequence 26, Appli
30	47	100.0	604	10	US-09-984-010-7	Sequence 7, Appli
31	47	100.0	608	14	US-10-165-603-24	Sequence 24, Appli
32	47	100.0	608	14	US-10-165-603-25	Sequence 25, Appli
33	47	100.0	608	14	US-10-316-253-42	Sequence 42, Appli
34	47	100.0	609	10	US-09-919-039-370	Sequence 370, App
35	47	100.0	609	12	US-10-609-346-12	Sequence 12, Appli
36	47	100.0	609	13	US-10-153-064-7	Sequence 7, Appli
37	47	100.0	609	14	US-10-153-604A-7	Sequence 7, Appli
38	47	100.0	609	14	US-10-365-623-23	Sequence 23, Appli
39	47	100.0	610	9	US-09-984-186-2	Sequence 2, Appli
40	47	100.0	610	14	US-10-237-657-2	Sequence 2, Appli
41	47	100.0	610	14	US-10-237-708-2	Sequence 2, Appli
42	47	100.0	610	14	US-10-237-866-2	Sequence 2, Appli
43	47	100.0	610	14	US-10-237-871-2	Sequence 2, Appli
44	47	100.0	610	14	US-10-237-624-2	Sequence 2, Appli
45	47	100.0	616	12	US-10-433-108-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-10-074-956-24
; Sequence 24, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-24

Query Match 100.0%; Score 47; DB 13; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 2
US-10-074-956-27
; Sequence 27, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956

```
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-27

Query Match          100.0%; Score 47; DB 13; Length 241;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
Db      116 AKQEPERNE 124

RESULT 3
US-10-074-956-28
; Sequence 28, Application US/10074956
; Publication No. US2002019332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08:91-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-28

Query Match          100.0%; Score 47; DB 13; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
Db      116 AKQEPERNE 124

RESULT 4
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medien & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001

; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBR1-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match          100.0%; Score 47; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 5
US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091585A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445

Query Match          100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
Db      92 AKQEPERNE 100

RESULT 6
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
```

STATE: DC
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,010
FILING DATE: 21-May-2002
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 09/091,873
FILING DATE: 25-JUN-1998
APPLICATION NUMBER: PCT/GB96/03164
FILING DATE: 19-DEC-1996
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 7
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF545
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 8
US-09-833-117-18

; Sequence 18, Application US/09833117
; Publication No. US20030171267A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadechi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match 100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 9
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BlyS)
; FILE REFERENCE: Dyx-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-932-322-445

Query Match 100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 10
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.

; APPLICANT: Sadeghi, Roma
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match 100.0%; Score 47; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 11
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match 100.0%; Score 47; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 12
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12

; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match 100.0%; Score 47; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 13
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for:
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match 100.0%; Score 47; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKQEPERNE 9
Db 92 AKQEPERNE 100

RESULT 14
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Xinglong Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 47; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
      |||||
Db      92 AKQEPERNE 100

RESULT 15
US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34

Query Match      100.0%; Score 47; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AKQEPERNE 9
      |||||
Db      92 AKQEPERNE 100

Search completed: April 19, 2004, 12:54:58
Job time : 3.78947 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 0.919668 Seconds
(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_247_252

Perfect score: 33

Sequence: 1 HGDLE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	109	1	US-08-448-196A-1
2	33	100.0	380	2	US-08-709-923-3
3	33	100.0	547	4	US-09-252-991A-27640
4	33	100.0	582	1	US-08-134-638-1
5	33	100.0	583	1	US-08-448-196A-4
6	33	100.0	583	1	US-08-448-196A-5
7	33	100.0	583	1	US-08-448-196A-6
8	33	100.0	584	1	US-08-448-196A-7
9	33	100.0	585	1	US-08-153-799-14
10	33	100.0	585	1	US-08-448-196A-3
11	33	100.0	585	2	US-08-984-176-1
12	33	100.0	585	2	US-08-702-572-2
13	33	100.0	585	3	US-08-769-746-2
14	33	100.0	585	4	US-10-153-064-5
15	33	100.0	609	1	US-08-223-619-3
16	33	100.0	609	1	US-08-433-037-4
17	33	100.0	609	4	US-08-897-956A-2
18	33	100.0	609	4	US-10-153-064-7
19	33	100.0	609	4	US-09-976-594-73
20	33	100.0	609	5	PCT-US95-04075-3
21	33	100.0	610	2	US-08-797-689-2
22	33	100.0	610	4	US-09-984-186-2
23	33	100.0	651	4	US-10-153-064-133
24	33	100.0	652	4	US-10-153-064-99
25	33	100.0	652	4	US-10-153-064-96
26	33	100.0	652	4	US-10-153-064-105
27	33	100.0	652	4	US-10-153-064-132

28	33	100.0	653	4	US-10-153-064-131	Sequence 131, Appl
29	33	100.0	656	4	US-10-153-064-130	Sequence 130, Appl
30	33	100.0	660	4	US-10-153-064-90	Sequence 90, Appl
31	33	100.0	660	4	US-10-153-064-93	Sequence 93, Appl
32	33	100.0	666	4	US-09-489-039A-12388	Sequence 12388, A
33	33	100.0	668	4	US-10-153-064-102	Sequence 102, Appl
34	33	100.0	676	4	US-10-153-064-95	Sequence 95, Appl
35	33	100.0	676	4	US-10-153-064-98	Sequence 98, Appl
36	33	100.0	676	4	US-10-153-064-104	Sequence 104, Appl
37	33	100.0	676	4	US-10-153-064-127	Sequence 127, Appl
38	33	100.0	676	4	US-10-153-064-129	Sequence 129, Appl
39	33	100.0	677	4	US-10-153-064-125	Sequence 125, Appl
40	33	100.0	680	4	US-10-153-064-123	Sequence 123, Appl
41	33	100.0	684	4	US-10-153-064-92	Sequence 92, Appl
42	33	100.0	692	4	US-10-153-064-101	Sequence 101, Appl
43	33	100.0	783	1	US-08-256-938-2	Sequence 2, Appl
44	33	100.0	787	1	US-08-256-938-4	Sequence 4, Appl
45	33	100.0	787	2	US-08-797-689-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-08-448-196A-1

; Sequence 1, Application US/08448196A

; Patent No. 5780594

; GENERAL INFORMATION:

; APPLICANT: CARTER, DANIEL C.

; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR RELATED PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NASA

; STREET: MARSHALL SPACE FLIGHT CENTER

; CITY: HUNTSVILLE

; STATE: ALABAMA

; COUNTRY: USA

; ZIP: 35812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/448,196A

; FILING DATE: 23-MAY-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: BROAD JR., ROBERT L.

; REGISTRATION NUMBER: 18,757

; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 205-544-0021

; TELEFAX: 205-544-0258

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 109 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: Protein

; FRAGMENT TYPE: Internal

US-08-448-196A-1

Query Match 100.0%; Score 33; DB 1; Length 109;

Best Local Similarity 100.0%; Pred. No. 8.9;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLE 6

Db 57 HGDLE 62

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27640

Query Match 100.0%; Score 33; DB 4; Length 547;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
Db 488 HGDLL 493

RESULT 4
US-08-134-638-1
; Sequence 1, Application US/08134638
; Patent No. 5473050
; GENERAL INFORMATION:
; APPLICANT: Strand, Frederick T
; TITLE OF INVENTION: Denatured Bovine Serum Albumin Milk
; TITLE OF INVENTION: Products and Method Therefor
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frederick T. Strand
; STREET: P.O. Box 64321
; CITY: Phoenix
; STATE: Arizona
; COUNTRY: USA
; ZIP: 85082-4321
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 1.2 Mb storage
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/134,638
; FILING DATE: 10/12/93
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiss, Harry M
; REGISTRATION NUMBER: 19,497
; REFERENCE/DOCKET NUMBER: 1795P1423
; TELEPHONE: (602) 994-8888
; TELEFAX: (602) 947-2683
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-134-638-1

Query Match 100.0%; Score 33; DB 1; Length 582;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
Db 245 HGDLL 250

RESULT 5
US-08-448-196A-4
; Sequence 4, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27640

Query Match 100.0%; Score 33; DB 4; Length 547;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
Db 488 HGDLL 493

RESULT 4
US-08-134-638-1
; Sequence 1, Application US/08134638
; Patent No. 5473050
; GENERAL INFORMATION:
; APPLICANT: Strand, Frederick T
; TITLE OF INVENTION: Denatured Bovine Serum Albumin Milk
; TITLE OF INVENTION: Products and Method Therefor
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frederick T. Strand
; STREET: P.O. Box 64321
; CITY: Phoenix
; STATE: Arizona
; COUNTRY: USA
; ZIP: 85082-4321
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 1.2 Mb storage
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/134,638
; FILING DATE: 10/12/93
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiss, Harry M
; REGISTRATION NUMBER: 19,497
; REFERENCE/DOCKET NUMBER: 1795P1423
; TELEPHONE: (602) 994-8888
; TELEFAX: (602) 947-2683
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-134-638-1

Query Match 100.0%; Score 33; DB 1; Length 582;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
Db 245 HGDLL 250

RESULT 5
US-08-448-196A-4
; Sequence 4, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

;; TITLE OF INVENTION: RELATED PROTEINS
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: NASA
;; STREET: MARSHALL SPACE FLIGHT CENTER
;; CITY: HUNTSVILLE
;; STATE: ALABAMA
;; COUNTRY: USA
;; ZIP: 35812
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; FILING DATE: 23-MAY-1995
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BROAD JR., ROBERT L.
;; REGISTRATION NUMBER: 18,757
;; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 205-544-0021
;; TELEFAX: 205-544-0258
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 583 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; US-08-448-196A-4

Query Match 100.0%; Score 33; DB 1; Length 583;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
Db 246 HGDLL 251

RESULT 6
US-08-448-196A-5
; Sequence 5, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.

;; REGISTRATION NUMBER: 18,757
;; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 205-544-0021
;; TELEFAX: 205-544-0258
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 583 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; US-08-448-196A-5

Query Match 100.0%; Score 33; DB 1; Length 583;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
Db 246 HGDLL 251

RESULT 7
US-08-448-196A-6
; Sequence 6, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-448-196A-6

Query Match 100.0%; Score 33; DB 1; Length 583;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
Db 246 HGDLL 251

RESULT 8
US-08-448-196A-7
; Sequence 7, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-7

Query Match 100.0%; Score 33; DB 1; Length 584;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
Db 247 HGDLL 252

RESULT 9
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA

QY 1 HGDLL 6
Db 247 HGDLL 252

RESULT 10
US-08-448-196A-3
; Sequence 3, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER

US-08-153-799-14

Query Match 100.0%; Score 33; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLL 6
Db 247 HGDLL 252

RESULT 10
US-08-448-196A-3
; Sequence 3, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER

```

; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-448-196A-3

Query Match 100.0%; Score 33; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
DB 247 HGDLLLE 252
|||||

RESULT 11
US-08-984-176-1
; Sequence 1, Application US/08984176
; Patent No. 5948609
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C
; APPLICANT: HO, JOSEPH X
; APPLICANT: RUKER, FLORIAN
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
; TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
; FILE REFERENCE: 08/984,176
; CURRENT APPLICATION NUMBER: US/08/984,176
; CURRENT FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1

Query Match 100.0%; Score 33; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
DB 247 HGDLLLE 252
|||||

RESULT 12
US-08-702-572-2
; Sequence 2, Application US/08702572
```

```

; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 33; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
DB 247 HGDLLLE 252
|||||

RESULT 13
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US/08/769,746
/ FILING DATE: 19-DEC-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Carroll, Peter G.
/ REGISTRATION NUMBER: 32,837
/ REFERENCE/DOCKET NUMBER: MBRI-02584
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 705-8410
/ TELEFAX: (415) 397-8338
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 595 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-769-746-2

Query Match          100.0%; Score 33; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
Db 247 HGDLLLE 252

RESULT 14
US-10-153-064-5
/ Sequence 5, Application US/10153064
/ Patent No. 6663485
/ GENERAL INFORMATION:
/ APPLICANT: Bell et al.
/ TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
/ FILE REFERENCE: PF556
/ CURRENT APPLICATION NUMBER: US/10/153,064
/ CURRENT FILING DATE: 2002-05-24
/ PRIOR APPLICATION NUMBER: 60/293,212
/ PRIOR FILING DATE: 2001-05-25
/ NUMBER OF SEQ ID NOS: 137
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 5
/ LENGTH: 585
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
/ US-10-153-064-5

Query Match          100.0%; Score 33; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
Db 247 HGDLLLE 252

RESULT 15
US-08-222-619-3
/ Sequence 3, Application US/08222619
/ Patent No. 5852352
/ GENERAL INFORMATION:
/ APPLICANT: Lichenstein, Henri
/ APPLICANT: Lyons, David
/ APPLICANT: Wurfel, Mark
/ APPLICANT: Wright, Samuel
/ TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
/ TITLE OF INVENTION: Protein
/ NUMBER OF SEQUENCES: 33
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Amgen Center, Patent Operations/RRC
/ STREET: 1840 DeHavilland Drive
/ CITY: Thousand Oaks
/ STATE: California

/ APPLICATION NUMBER: US/08/222,619
/ FILING DATE:
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/222,619
/ FILING DATE:
/ CLASSIFICATION: 435
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 609 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ US-08-222-619-3

Query Match          100.0%; Score 33; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
Db 271 HGDLLLE 276

Search completed: April 19, 2004, 12:05:20
Job time : 1.91967 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.4626 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_266_277

Perfect score: 57

Sequence: 1 ENQDSISSKLKE 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: Pirl:*

2: Pirl2:*

3: Pirl3:*

4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	57	100.0	600	2 A47391	serum albumin prec
2	57	100.0	609	1 ABHUS	serum albumin prec
3	54	94.7	265	2 I45986	albumin - dog (fra
4	54	94.7	608	2 S57632	serum albumin prec
5	51	89.5	605	1 ABFGS	serum albumin prec
6	51	89.5	607	1 ABROS	serum albumin prec
7	45	78.9	423	2 S41122	mannose-6-phosphat
8	44	77.2	599	1 A54906	afamin precursor -
9	44	77.2	607	1 ABHS	serum albumin prec
10	43	75.4	607	1 ABROS	serum albumin prec
11	40	70.2	609	2 JCS838	albumin - Mongolia
12	40	70.2	2469	2 H36812	hypothetical prote
13	39	68.4	608	2 A53195	afamin precursor -
14	38	66.7	238	2 T32033	hypothetical prote
15	38	66.7	1091	2 T34107	hypothetical prote
16	37	64.9	416	2 P68640	hippurate hydratase
17	37	64.9	453	2 A05139	serum albumin - mo
18	37	64.9	495	2 F82123	probable flagellar
19	37	64.9	535	2 S31097	cold acclimation p
20	37	64.9	556	2 B82882	conserved hypothet
21	37	64.9	563	2 AH2975	hypothetical prote
22	37	64.9	563	2 C98307	opha protein (impo
23	37	64.9	581	2 F84857	hypothetical prote
24	37	64.9	608	1 ABHUS	serum albumin prec
25	36	63.2	176	2 F81346	heat shock protein
26	36	63.2	208	2 B30305	submandibular glan
27	36	63.2	282	2 T24693	hypothetical prote
28	36	63.2	298	2 G75140	hypothetical prote
29	36	63.2	419	2 T44276	transducer protein

Htr12 transducer [i
protein R02C3.4 [i
preprotein translo
protein F10A5.15 [i
probable finger pr
large tegument pro
heavy metal bindin
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
conserved hypothet
B. subtilis YgfA p
B. subtilis YgfA p
merozoite 44K vari
farnesyl-diphospha

30 36 63.2 420 2 C84298
31 36 63.2 575 2 C88924
32 36 63.2 1011 2 S65668
33 36 63.2 1037 2 D96786
34 36 63.2 1170 2 S30010
35 36 63.2 2471 2 T42977
36 35 61.4 67 2 G72392
37 35 61.4 79 2 B97761
38 35 61.4 90 2 T44137
39 35 61.4 248 2 A69173
40 35 61.4 264 2 T33614
41 35 61.4 296 1 G64453
42 35 61.4 314 2 AC1484
43 35 61.4 314 2 A11123
44 35 61.4 317 2 A48443
45 35 61.4 416 2 S52075

ALIGNMENTS

RESULT 1

A47391
serum albumin precursor - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C;Accession: A47391
R;Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A;Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilir
A;Reference number: A47391; MUID:93211971; PMID:8460152
A;Contents: B/B homozygote
A;Accession: A47391
A;Status: preliminary
A;Molecule type: mRNA; protein
A;Residues: 1-600 <WAT>
A;Cross-references: GB:M90463; NID:g342294; PIDN:AAA36906.1; PID:g342295
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBI:128280, NCBI:128281)
C;Superfamily: serum albumin; serum albumin repeat homology
F;21-194/Domain: serum albumin repeat homology <SA1>
F;213-386/Domain: serum albumin repeat homology <SA2>
F;405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 57; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 0.009;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ENQDSISSKLKE 12
|||
Db 282 ENQDSISSKLKE 293

RESULT 2

ABHUS
serum albumin precursor [validated] - human
N;Alternate names: preproalbumin
N;Contains: kinetensin
C;Species: Homo sapiens (man)
C;Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text_change 17-Mar-2000
C;Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S3
R;Law, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu
Nucleic Acids Res. 9, 6103-6114, 1981
A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia col
A;Reference number: A93743; MUID:82081882; PMID:6171778
A;Accession: A93743
A;Molecule type: mRNA
A;Residues: 1-419, 'K', 421-609 <LAW>
A;Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:g28591; PIDN:CA2
R;Dugaiczky, A.; Law, S.W.; Dennison, O.E
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A;Reference number: A93936; MUID:82105994; PMID:6275391

A;Accession: A93936
A;Molecule type: mRNA
A;Residues: 1-120,'G',122-609 <DUG>
A;Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590
R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A;Reference number: 139427; MUID:86140099; PMID:2419329
A;Accession: I39427
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-26 <RA>
A;Cross-references: GB:MI3075; NID:G178330; PIDN:AAA51588.1; PID:G553173
R;Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.
A;Reference number: 159286; MUID:94181575; PMID:9134387
A;Accession: I59286
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 282-290, 'KSRPDLQ' <WAT>
A;Cross-references: GB:IS69192; NID:G545032; PIDN:AA30282.1; PID:G546033
A;Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia
R;Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-
A;Reference number: 159313; MUID:94294404; PMID:8022807
A;Accession: I59313
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 589-590, 'ALPRPVKLLQVLP' <MAD>
A;Cross-references: GB:S70799; NID:G547231; PIDN:AA31177.1; PID:G547232
A;Note: This frame-shift variant is designated albumin Bazzano; four additional variants
R;Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A;Reference number: G08292
A;Accession: G01747
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-120,'G',122-455 <MEN>
A;Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431
R;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
A;Reference number: S55314; MUID:95275251; PMID:7755581
A;Accession: S55314
A;Molecule type: protein
A;Residues: 19-27 <LED>
R;Meloun, B.; Moravek, L.; Koetka, V.
FEBS Lett. 58, 134-137, 1975
A;Title: Complete amino acid sequence of human serum albumin.
A;Reference number: A91420; MUID:76187907; PMID:1225573
A;Accession: A91420
A;Molecule type: protein
A;Residues: 25-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',395-
R;Coehr, U.; Spitteller, G.; Tripiet, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from ur
A;Reference number: S06422
A;Accession: S06422
A;Molecule type: protein
A;Residues: 25-48 <ROE>
R;Pinch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A;Title: Mass spectrometric identification of modifications to human serum albumin treat
A;Reference number: S36882; MUID:93384321; PMID:8373198
A;Accession: S36882
A;Molecule type: protein
A;Residues: 45-67,'141-160,'311-337,'469-490,'570-581 <PIN>
R;Kausler, E.; Spitteller, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A;Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol

A;Reference number: S17599; MUID:92126241; PMID:1772598
A;Accession: S17599
A;Molecule type: protein
A;Residues: 25-54,'354-357,'431-447 <KAU>
A;Note: 49-Leu was also found
R;Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid protea
A;Reference number: A03239; MUID:86242180; PMID:3087352
A;Accession: A03239
A;Molecule type: protein
A;Residues: 166-173,'L' <MOG>
R;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S.
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A;Title: Mutations in genetic variants of human serum albumin found in Italy.
A;Reference number: A38255; MUID:91062352; PMID:2247440
A;Accession: C38255
A;Molecule type: protein
A;Residues: 76-83,'K',85-106 <GAL3>
A;Note: this variant is designated albumin Torino
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin He
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: protein
A;Residues: 255-263,'E',265-281 <MIN1>
A;Note: this variant is designated albumin Herborn
R;Minchiotti, L.; Galliano, M.; Scoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta
Biochim. Biophys. Acta 1119, 232-238, 1992
A;Title: Two albumins with identical electrophoretic mobility are produced by diffe
A;Reference number: S21078; MUID:92190239; PMID:1347703
A;Accession: S21078
A;Molecule type: protein
A;Residues: 354-356,'K',358-378 <MIN2>
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported
R;He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A;Title: Atomic structure and chemistry of human serum albumin.
A;Reference number: A46756; MUID:92334427; PMID:1630489
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.
in: The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40
A;Reference number: A94442
A;Contents: annotation; three-dimensional structure and disulfide bonds
R;Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A;Title: Disulfide bonds in human serum albumin.
A;Reference number: A90930
A;Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A;Reference number: A90299; MUID:78186630; PMID:656055
A;Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
in: Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20
A;Title: Serum albumin: conformation and active sites.
A;Reference number: A94408

A;Contents: annotation; binding sites
R;Harper, M.E.; Dugaiczky, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes
A;Reference number: A90028; MUID:8327982; PMID:6192711
A;Contents: annotation; gene position
R;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.
A;Reference number: A46755; MUID:76257808; PMID:955075
A;Contents: annotation
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid)
R;Bohney, J.P.; Fonda, M.L.; Feidhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A;Title: Identification of lys(190) as the primary binding site for pyridoxal 5'-phosphate
A;Reference number: A56294; MUID:92183881; PMID:1544460
A;Contents: annotation
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in B
atase activity
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
liver, protophyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
C;Comment: A large number of variants of human serum albumin have been described.
C;Genetics:
A;Gene: GDB:ALB
A;Cross-references: GDB:118990; OMIM:103600
A;Map position: 4q11-4q13
C;Superfamily: serum albumin; serum albumin repeat homology
F;1-18/Domain: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox
F;19-24/Domain: signal sequence #status predicted <SIG>
F;25-609/Product: propeptide #status experimental <PRO>
F;29-202/Domain: serum albumin #status experimental <PRO>
F;166-174/Product: kinetensin #status experimental <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,148-193,192-203,224-270,269-277,289-303,302-313,340-385,394-393,4
F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 57; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.0092;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12
|||||:|||||
DB 290 ENQDSISSKLKE 301

RESULT 3
146986
albumin - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 20-Aug-1999
C;Accession: I46986
R;Spitzauer, S.; Schweiger, C.; Sperr, W.R.; Pandjaitan, B.; Valent, P.; Muhl, S.; Ebner
J. Allergy Clin. Immunol. 93, 614-627, 1994
A;Title: Molecular characterization of dog albumin as a cross-reactive allergen.
A;Reference number: I46986; MUID:94204492; PMID:7512102
A;Accession: I46986
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-265 <SPI>
A;Cross-references: GB:S72946; NID:9633937; PIDN:AAB30434.1; PID:g633938
C;Superfamily: serum albumin; serum albumin repeat homology
F;7-180/Domain: serum albumin repeat homology <SA2>

Query Match 94.7%; Score 54; DB 2; Length 265;
Best Local Similarity 91.7%; Pred. No. 0.014;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12
|||||:|||||
DB 76 ENQDSISSKLKE 87

RESULT 4
S57632
serum albumin precursor - cat
C;Species: Felis silvestris catus (domestic cat)
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
C;Accession: JC4660; S57632
R;Hilger, C.; Grigioni, F.; Hentges, F.
Gene 169, 295-296, 1996
A;Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A;Reference number: JC4660; MUID:96194924; PMID:8647469
A;Accession: JC4660
A;Molecule type: mRNA
A;Residues: 1-608 <HI2>
A;Cross-references: EMBL:X84842; NID:9886484; PIDN:CAA59279.1; PID:g886485
A;Experimental source: liver
C;Comment: This protein is the major protein component in plasma. It functions as a multi
ein has 35 conserved cysteine residues.
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: liver; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRP>
F;25-608/Product: serum albumin #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>

Query Match 94.7%; Score 54; DB 2; Length 608;
Best Local Similarity 91.7%; Pred. No. 0.034;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12
|||||:|||||
DB 290 ENQDSISSKLKE 301

RESULT 5
AB03
serum albumin precursor - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-Jun-1999
C;Accession: S01382; A61006
R;Weinstock, J.; Baldwin, G.S.
Nucleic Acids Res. 16, 9045, 1988
A;Title: Nucleotide sequence of porcine liver albumin.
A;Reference number: S01382; MUID:89016582; PMID:3174440
A;Accession: S01382
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-605 <WE1>
A;Cross-references: EMBL:X12422; NID:91875; PIDN:CAA30970.1; PID:g833798
R;Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
J. Bone Miner. Res. 4, 235-241, 1989
A;Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral
A;Reference number: A61006; MUID:89269769; PMID:2728927
A;Accession: A61006
A;Molecule type: protein
A;Residues: 23-51, 'X', 53-54; 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XV', 155 <LIM>
A;Experimental source: dental enamel
A;Note: albumin and other serum proteins are also found in bone
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-16/Domain: signal sequence (fragment) #status predicted <PRO>
F;17-22/Domain: propeptide #status predicted <PRP>
F;23-605/Product: serum albumin #status predicted <MAT>
F;27-199/Domain: serum albumin repeat homology <SA1>
F;218-391/Domain: serum albumin repeat homology <SA2>
F;410-589/Domain: serum albumin repeat homology <SA3>
F;76-84, 97-113, 112-123, 145-190, 189-198, 221-267, 286-294, 286-300, 299-310, 337-382, 381-390, 4
F;261/Binding site: bilirubin (Lys) #status predicted

Query Match 89.5%; Score 51; DB 1; Length 605;
Best Local Similarity 83.3%; Pred. No. 0.13;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12
|||||:|||||
Db 287 ENQDITSLKE 298

RESULT 6
ABOS
serum albumin precursor [validated] - bovine
N/Alternate names: 67K protein; preproalbumin
C/Species: Bos primigenius taurus (cattle)
C/Date: 24-Apr-1984 #sequence revision 30-Sep-1993 #text change 18-Aug-2000
C/Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94
R/Holowachuk, E.W.; Stoltenberg, J.K.; Reed, R.G.; Peters Jr., T.
submitted to the EMBL Data Library, August 1991
A/Description: Bovine serum albumin: cDNA sequence and expression.
A/Reference number: A38885
A/Accession: A38885
A/Molecule type: mRNA
A/Residues: 1-607 <HOL>
A/Cross-references: EMBL:W73215
R/Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
A/Title: Rapid confirmation and revision of the primary structure of bovine serum albumin
A/Reference number: A36401; MUID:91083649; PMID:2260975
A/Accession: A36401
A/Molecule type: protein
A/Residues: 25-41, H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <Hir>
R/MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-485, 1979
A/Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A/Reference number: A91258; MUID:80024578; PMID:488109
A/Accession: A91258
A/Molecule type: protein
A/Residues: 1-32 <MAG>
R/Hsieh, J.C.; Lin, F.P.; Tam, M.F.
Anal. Biochem. 170, 1-8, 1988
A/Title: Electrophoretic onto glass-fiber filter from an analytical isoelectrofocusing g
A/Reference number: A60808; MUID:88267456; PMID:3389500
A/Accession: B60808
A/Molecule type: protein
A/Residues: 25-41 <HSI>
R/Strawich, E.; Glincher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A/Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelins' is albu
A/Reference number: S10780; MUID:90336641; PMID:2379503
A/Accession: S10780
A/Molecule type: protein
A/Residues: 25-41, H', 43-57, 59-64 <STR>
R/Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A/Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A/Reference number: A45800; MUID:89341406; PMID:2474609
A/Accession: D45800
A/Molecule type: protein
A/Residues: 163-172 <CAR>
R/Carraway, R.E.; Mitra, S.P.; Cochrane, D.E.
J. Biol. Chem. 262, 5968-5973, 1987
A/Title: Structure of a biologically active neurotensin-related peptide obtained from pe
A/Reference number: A26693; MUID:87194805; PMID:2437111
A/Accession: A26693
A/Molecule type: protein
A/Residues: 165-172, 'L', <CA2>
R/Reed, R.G.; Putnam, F.W.; Peters Jr., T.
Biochem. J. 191, 867-868, 1980
A/Title: Sequence of residues 400-403 of bovine serum albumin.
A/Reference number: A90309; MUID:82023364; PMID:7283978
A/Accession: A90309
A/Molecule type: protein

A/Residues: 402-433 <REE>
R/Brown, J.R.
Fed. Proc. 34, 591, 1975
A/Title: Structure of bovine serum albumin.
A/Reference number: A91458
A/Accession: A91458
A/Molecule type: protein
A/Residues: 25-41, H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'I'
R/Brown, J.R.
submitted to the Atlas, April 1975
A/Reference number: A94551
A/Accession: A94551
A/Molecule type: protein
A/Residues: 190-195 <BR2>
R/Brown, J.R.
Fed. Proc. 33, 1389, 1974
A/Reference number: A91457
A/Contents: annotation; disulfide bonds
R/Werlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A/Title: Preparation and characterization of novel substrates of insulin proteinase (EC
A/Reference number: S55232; MUID:95031935; PMID:7945219
A/Accession: S55232
A/Status: preliminary
A/Molecule type: protein
A/Residues: 529-536; 569-572 <WER>
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: carrier protein; copper binding; duplication; plasma
F:1-18/Domain: signal sequence #status experimental <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:23-607/Product: serum albumin #status experimental <MPT>
F:23-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:21/Binding site: copper (His) #status predicted
F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392, 'I'

Query Match 89.5%; Score 51; DB 1; Length 607;
Best Local Similarity 83.3%; Pred. No. 0.13;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12
|||||:|||||
Db 289 DNQDTISSKLKE 300

RESULT 7
S41122
mannose-6-phosphate isomerase (EC 5.3.1.8) - human
N/Alternate names: phosphomannose isomerase
C/Species: Homo sapiens (man)
C/Date: 07-Sep-1994 #sequence revision 10-Nov-1995 #text_change 08-Oct-1999
C/Accession: S41122; S38666
R/Proudfoot, A.E.I.; Turcatti, G.; Wells, T.N.C.; Payton, M.A.; Smith, D.J.
Eur. J. Biochem. 219, 415-423, 1994
A/Title: Purification, cDNA cloning and heterologous expression of human phosphomannose
A/Reference number: S41122; MUID:94139717; PMID:8307007
A/Accession: S41122
A/Molecule type: mRNA
A/Residues: 1-423 <PRO>
A/Cross-references: EMBL:X76057; NID:g416016; PIDN:CAAS3657.1; PID:g416017
C/Genetics:
A/Gene: GDB:MPI
A/Cross-references: GDB:119397; OMIM:154550
A/Map position: 15q22-15qter
C/Superfamily: yeast mannose-6-phosphate isomerase
C/Keywords: intramolecular oxidoreductase; isomerase

Query Match 78.9%; Score 45; DB 2; Length 423;
Best Local Similarity 66.7%; Pred. No. 1.2;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12

Db 77 ENQDSLGSVKVD 88
|||||:|:|:|

RESULT 8
A:Accession: A54906
A:Alternate names: alpha-albumin
N:Alternate names: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Mar-2000
C:Accession: A54906; JG6143; S68554; S78082; I39425
R:Lichenstein, H.S.; Lyons, D.E.; Wurfel, M.M.; Johnson, D.A.; McGinley, M.D.; Leidli, J.
J. Biol. Chem. 269, 18149-18154, 1994
A:Title: Afamin is a new member of the albumin, alpha-fetoprotein, and vitamin D-binding
A:Reference number: A54906; MUID:94299534; PMID:7517938
A:Accession: A54906
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-599 <LIC>
A:Cross-references: GB:L32140; NID:G533885; PIDN:AAA21612.1; PID:G547402
R:Mishio, H.; Dugaiczkyk, A.
Proc. Natl. Acad. Sci. U.S.A. 93, 7557-7561, 1996
A:Title: Complete structure of the human alpha-albumin gene, a new member of the serum
A:Reference number: JG6143; MUID:96353855; PMID:8755513
A:Accession: JG6143
A:Molecule type: DNA
A:Residues: 1-599 <NIS>
A:Cross-references: GB:U51243; NID:G1418261; PIDN:AAC50720.1; PID:G1418262
R:Nishio, H.; Heiskanen, M.; Palotie, A.; Belanger, L.; Dugaiczkyk, A.
J. Mol. Biol. 259, 113-119, 1996
A:Title: Tandem arrangement of the human serum albumin multigene family in the sub-centr
A:Reference number: S68554; MUID:96240683; PMID:8648639
A:Accession: S68554
A:Molecule type: DNA
A:Residues: 1-29 <NIW>
A:Cross-references: GB:U51243; NID:G1418261
A:Note: neither the complete nucleic acid sequence nor the complete translation are show
R:Nishio, H.; Heiskanen, M.; Palotie, A.; Belanger, L.; Dugaiczkyk, A.
submitted to the EMBL Data Library, March 1996
A:Description: Tandem arrangement of the human serum albumin multigene family in the sub
A:Reference number: S78082
A:Accession: S78082
A:Molecule type: DNA
A:Residues: 1-599 <NIS>
A:Cross-references: EMBL:U51243; NID:G1418261; PIDN:AAC50720.1; PID:G1418262
R:Allard, D.; Gilbert, S.; Lamontagne, A.; Hamel, D.; Belanger, L.
Gene 153, 287-288, 1995
A:Title: Identification of rat alpha-albumin and cDNA cloning of its human ortholog.
A:Reference number: I39424; MUID:96180738; PMID:7875606
A:Accession: I39425
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 105-207 <ALL>
A:Cross-references: GB:L35497; NID:G530131; PIDN:AAA68198.1; PID:G857676
C:Gene: GDB:AFM
A:Gene: GDB:376475; OMIM:104145
A:Map position: 4q11-q13
A:Introns: 30/1, 46/2, 90/3, 161/2, 205/3, 238/2, 281/3, 353/2, 397/3, 430/2, 474/3, 549
C:Superfamily: serum albumin; serum albumin repeat homology
A:Keywords: extracellular protein; glycoprotein; plasma
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-599/Product: afamin #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-590/Domain: serum albumin repeat homology <SA3>
F:33,109,402,488/Binding site: carbohydrate (Asn) #status predicted

Query Match 77.2%; Score 44; DB 1; Length 599;
Best Local Similarity 90.0%; Pred. No. 2.7;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDSISSKLKE 12

Db 292 QDSISSKLKE 301
|||||:|:|:|

RESULT 9
ABSHS
serum albumin precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S06936
R:Brown, W.M.; Dziugielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A:Reference number: S06936; MUID:90098888; PMID:2602160
A:Accession: S06936
A:Molecule type: mRNA
A:Residues: 1-607 <BRO>
A:Cross-references: EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membr
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 77.2%; Score 44; DB 1; Length 607;
Best Local Similarity 66.7%; Pred. No. 2.8;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLKE 12

Db 289 DHQDALSXLKE 300

RESULT 10

ABHOS
serum albumin precursor - horse
C:Species: Equus caballus (domestic horse)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S34053
R:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A:Reference number: S34053; MUID:93345495; PMID:8344282
A:Accession: S34053
A:Molecule type: mRNA
A:Residues: 1-607 <HOA>
A:Cross-references: GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membr
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 75.4%; Score 43; DB 1; Length 607;
Best Local Similarity 81.8%; Pred. No. 4.3;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKXK 11
 :|||||
 Db 289 ENQDSISSKXK 299

RESULT 11

UC5838
 albumin - Mongolian jird
 C:Species: Meriones unguiculatus (Mongolian jird)
 C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000
 C:Accession: UC5838
 R:Yoshida, K.; Seto-Ohshtima, A.; Sanochara, H.
 DNA Res. 4, 351-354, 1997
 A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in the
 A:Reference number: UC5838; MUID:9811663; PMID:9455485
 A:Accession: UC5838
 A:Molecule type: mRNA
 A:Residues: 1-609 <YOS>
 A:Cross-references: DBJ:AB006197; NID:g2317277; PIDN:BAA21765.1; PID:g2317278
 A:Experimental source: liver
 C:Superfamily: serum albumin; serum albumin repeat homology
 F:222-395/Domain: serum albumin repeat homology <SA2>

Query Match 70.2%; Score 40; DB 2; Length 609;
 Best Local Similarity 81.8%; Pred. No. 16;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKXK 11
 :|||||
 Db 291 ENQDSISSKXK 301

RESULT 12

H36812
 hypothetical protein ORF64 - saimirine herpesvirus 1 (strain 11)
 C:Species: saimirine herpesvirus 1
 A:Note: host Saimiri sciureus (common squirrel monkey)
 C:Date: 16-Oct-1992 #sequence_revision 16-Oct-1992 #text_change 08-Oct-1999
 C:Accession: H36812
 R:Albrecht, J.
 submitted to the EMBL Data Library, January 1992
 A:Description: Primary structure of the herpesvirus saimir genome.
 A:Reference number: A36806
 A:Accession: H36812
 A:Molecule type: DNA
 A:Residues: 1-2469 <ALB>
 A:Cross-references: GB:X64346; NID:g60320; PIDN:CAA45687.1; PID:g60385
 R:Albrecht, J. C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; Wi
 J. Virol. 66, 5047-5059, 1992
 A:Title: Primary structure of the herpesvirus saimir genome.
 A:Reference number: A37309; MUID:92333688; PMID:1321287
 A:Contents: annotation; protein-coding frames
 A:Note: neither protein nor nucleotide sequence is given
 C:Genetics:
 A:Gene: 64

Query Match 70.2%; Score 40; DB 2; Length 2469;
 Best Local Similarity 58.3%; Pred. No. 75;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKXKE 12
 :|||||
 Db 373 EDDDNVTSKXKE 384

RESULT 13

A53195
 afamin precursor - rat
 N:Alternate names: alpha-albumin
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-2000
 C:Accession: A53195
 R:Belanger, L.; Roy, S.; Allard, D.

J. Biol. Chem. 269, 5481-5484, 1994
 A:Title: New albumin gene 3', adjacent to the alpha-1-fetoprotein locus.
 A:Reference number: A53195; MUID:94164881; PMID:7509788
 A:Accession: A53195
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-608 <BEL>
 A:Cross-references: GB:X76456; NID:g456358; PIDN:CAA53994.1; PID:g456359
 C:Genetics:
 A:Introns: 30/1; 47/2; 91/3; 162/2
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: glycoprotein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-608/Product: afamin #status predicted <NAT>
 F:29-202/Domain: serum albumin repeat homology <SA1>
 F:221-394/Domain: serum albumin repeat homology <SA2>
 F:413-590/Domain: serum albumin repeat homology <SA3>

Query Match 68.4%; Score 39; DB 2; Length 608;
 Best Local Similarity 88.9%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QDSISSKXK 11
 :|||||
 Db 292 QDSISSKXK 300

RESULT 14

T32033
 hypothetical protein F41B5.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T32033
 R:Dante, M.; Kramer, J.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of C. elegans cosmid F41B5.
 A:Reference number: Z21115
 A:Accession: T32033
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-238 <DAN>
 A:Cross-references: EMBL:AF016676; PIDN:AAC25902.1; GSPDB:GN00023; CESP:F41B5.5
 A:Experimental source: strain Bristol N2; clone F41B5
 C:Genetics:
 A:Gene: CESP:F41B5.5
 A:Map position: 5
 A:Introns: 77/1; 132/1

Query Match 66.7%; Score 38; DB 2; Length 238;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ENQDSISSKXKE 12
 :|||||
 Db 163 ENQDSABAYLKE 174

RESULT 15

T34107
 hypothetical protein C18C4.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34107
 R:Gattung, S.
 submitted to the EMBL Data Library, April 1996
 A:Description: The sequence of C. elegans cosmid C18C4.
 A:Reference number: Z21478
 A:Accession: T34107
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-1091 <GAT>
 A:Cross-references: EMBL:U55369; PIDN:AAC25825.1; GSPDB:GN00023; CESP:C18C4.5
 A:Experimental source: strain Bristol N2; clone C18C4

C:Genetics:
 A:Gene: CESP:C18C4.5
 A:Map position: 5
 A:Introns: 38/3; 95/3; 179/2; 259/3; 301/3; 419/2; 573/3; 613/1; 875/3; 920/2; 959/2; 98

Query Match 66.7%; Score 38; DB 2; Length 1091;
 Best Local Similarity 58.3%; Pred. No. 74;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ENQDSISSKLYE 12
 :|||:|
 Db 189 QNMDSLSEKLS 200

Search completed: April 19, 2004, 12:02:28
 Job time : 2.4626 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.875346 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_266_277

Perfect score: 57
Sequence: 1 ENQDSISSKLKE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	600	1	ALBU_MACMU
2	57	100.0	609	1	ALBU_HUMAN
3	54	94.7	608	1	ALBU_CANFA
4	54	94.7	608	1	ALBU_FELCA
5	51	89.5	605	1	ALBU_PIG
6	51	89.5	607	1	ALBU_BOVIN
7	45	78.9	422	1	MANA_HUMAN
8	44	77.2	599	1	AFAM_HUMAN
9	44	77.2	607	1	ALBU_SHEEP
10	43	75.4	607	1	ALBU_HORSE
11	40	70.2	609	1	ALBU_MERUN
12	40	70.2	2469	1	TEGU_HSVSA
13	39	68.4	608	1	AFAM_RAT
14	39	68.4	608	1	ALBU_RABIT
15	39	68.4	611	1	AFAM_MOUSE
16	37	64.9	416	1	YTNL_BACSU
17	37	64.9	608	1	ALBU_MOUSE
18	37	64.9	608	1	ALBU_RAT
19	36	63.2	176	1	GREP_CAMJZ
20	36	63.2	1011	1	SECA_PEA
21	36	63.2	1170	1	YKDB_YEAST
22	36	63.2	1679	1	GC22_MOUSE
23	35	61.4	248	1	PCRB_MEITH
24	35	61.4	296	1	YC32_METUA
25	35	61.4	416	1	FDFT_MOUSE
26	35	61.4	417	1	FDFT_HUMAN
27	35	61.4	504	1	SIX1_YEAST
28	35	61.4	626	1	CC23_YEAST
29	35	61.4	635	1	DNAK_BORBU
30	35	61.4	637	1	SYT_GLOB
31	35	61.4	676	1	SPL1_COTUA
32	35	61.4	852	1	RA50_THEME
33	35	61.4	1312	1	L159_CAEEL

34 35 61.4 1324 1 SMC4_SCHPO
35 35 61.4 3680 1 DMD_CANFA
36 35 61.4 3685 1 DMD_HUMAN
37 34 59.6 183 1 ANF2_XENLA
38 34 59.6 187 1 ANF1_XENLA
39 34 59.6 217 1 EXPI_ERWCA
40 34 59.6 267 1 RECK_STABP
41 34 59.6 354 1 HOM1_HUMAN
42 34 59.6 366 1 HOM1_MOUSE
43 34 59.6 366 1 HOM1_RAT
44 34 59.6 403 1 YME8_YEAST
45 34 59.6 439 1 ENGA_THEME

ALIGNMENTS

RESULT 1
ALBU_MACMU STANDARD; PRT; 600 AA.
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93211971; PubMed=8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwulet J., Putnam F.W.;
RT "cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding.";
RL Proc Natl Acad Sci U S A. 90:2409-2413(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M90463; AAA36906.1; -.
CC PIR; A47391; A47391.
CC HGSP; P02769; ILE7B.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
CC NON TER 1
CC SIGNAL <1 10 BY SIMILARITY.
CC PROPEP 11 16 BY SIMILARITY.
CC CHAIN 17 600 SERUM ALBUMIN.
CC DOMAIN 17 197 ALBUMIN 1.
CC DOMAIN 204 389 ALBUMIN 2.
CC DOMAIN 396 587 ALBUMIN 3.
CC FT

P41004 schizosacch
O97592 canis famil
P11532 homo sapien
Q91617 xenopus lae
Q91898 xenopus lae
P33882 erwinia car
Q8crv9 staphylococ
Q86ym7 homo sapien
Q92zy3 mus musculu
Q92z14 rattus norv
Q04697 saccharomyc
Q9x1f8 thermotoga

FT METAL 19 19 COPPER (BY SIMILARITY).
 FT BINDING 256 256 BILIRUBIN (POTENTIAL).
 FT DISULFID 69 78 BY SIMILARITY.
 FT DISULFID 91 107 BY SIMILARITY.
 FT DISULFID 106 117 BY SIMILARITY.
 FT DISULFID 140 185 BY SIMILARITY.
 FT DISULFID 184 193 BY SIMILARITY.
 FT DISULFID 216 262 BY SIMILARITY.
 FT DISULFID 261 269 BY SIMILARITY.
 FT DISULFID 281 295 BY SIMILARITY.
 FT DISULFID 294 305 BY SIMILARITY.
 FT DISULFID 332 377 BY SIMILARITY.
 FT DISULFID 376 385 BY SIMILARITY.
 FT DISULFID 408 454 BY SIMILARITY.
 FT DISULFID 453 464 BY SIMILARITY.
 FT DISULFID 477 493 BY SIMILARITY.
 FT DISULFID 492 503 BY SIMILARITY.
 FT DISULFID 530 575 BY SIMILARITY.
 FT DISULFID 574 583 BY SIMILARITY.
 SQ SEQUENCE 600 AA; 67880 MW; B45C871A670E740B CRC64;
 Query Match 100.0%; Score 57; DB 1; Length 600;
 Best Local Similarity 100.0%; Pred. No. 0.0078;
 Matches 1; Conservative 0; Mismatches 0; Gaps 0; Indels 0;
 Oy 1 ENODSISSKLE 12
 Db 282 ENODSISSKLE 293
 RESULT 2
 ID ALBU_HUMAN STANDARD; PRT; 609 AA.
 AC P02768; Q95574; Q13140; Q9P157; Q9P117; Q9UH53; Q9UJZ0;
 DT 21-JUL-1996 (Rel. 01, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DE 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Serum albumin precursor.
 GN ALB
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86196112; PubMed=3009475;
 RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
 RA Beattie W.G., Dugaiczky A.,
 RT "Molecular structure of the human albumin gene is revealed by
 RT nucleotide sequence within q11-22 of chromosome 4.";
 RL J. Biol. Chem. 261:6747-6757(1986).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
 RX MEDLINE=82081882; PubMed=6171778;
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
 RA Najarian R.C., Seeburg P.H., Wion K.L.,
 RT "The sequence of human serum albumin cDNA and its expression in E.
 RT coli.";
 RL Nucleic Acids Res. 9:6103-6114(1981).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
 RX MEDLINE=82105994; PubMed=6275391;
 RA Dugaiczky A., Law S.W., Dennison O.E.,
 RT "Nucleotide sequence and the encoded amino acids of human serum
 RT albumin mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
 RN [4]
 RP SEQUENCE FROM N.A.
 TISSUE=Liver;
 RC Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
 RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (PROC903/PRO1708/PRO2044/PRO2619/PRO2675).

RC TISSUE=Petal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
 RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
 RT "Functional prediction of the coding sequences of 121 new genes
 RT deduced by analysis of cDNA clones from human fetal liver.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
 RA Huang M.C., Wu H.T.;
 RT "The cDNA sequences of human serum albumin.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 TISSUE=Liver, and Skeletal muscle;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Ziegler B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Spapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SEQUENCE OF 25-609.
 RX MEDLINE=76197907; PubMed=1225573;
 RA Meloun B., Moravsek L., Kostka V.;
 RT "Complete amino acid sequence of human serum albumin.";
 RL FEBS Lett. 58:134-137(1975).
 RN [9]
 RP SEQUENCE OF 25-609.
 RA Brown J.R., Shockley P., Behrens P.Q.;
 RL (In) Bing D.H. (eds.);
 RT The chemistry and physiology of the human plasma proteins, pp.23-40,
 RN Pergamon Press, New York (1979).
 RP SEQUENCE OF 1-455 FROM N.A.
 RC TISSUE=Liver;
 RA Menaya J., Parrilla R., Ayuso M.S.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE=86140099; PubMed=2419329;
 RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
 RT "The human albumin gene. Characterization of the 5' and 3' flanking
 RT regions and the polymorphic gene transcripts.";
 RL J. Biol. Chem. 261:3244-3251(1986).
 RN [12]
 RP SEQUENCE OF 222-229.
 RX MEDLINE=76257808; PubMed=955075;
 RA Walker J.E.;
 RT "Lysine residue 199 of human serum albumin is modified by
 RT acetylsalicylic acid.";
 RL FEBS Lett. 66:173-175(1976).
 RN [13]
 RP SEQUENCE OF 25-44 AND 480-499.
 RC TISSUE=Heart;
 RX MEDLINE=95203287; PubMed=7895732;
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 RT 1994."

Electrophoresis 15:1459-1465 (1994).
[14] DISULFIDE BONDS.
RA Saber M.A., Stockbauer P., Moravsek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579 (1977).
[15] BIILIRUBIN-BINDING SITE.
RA Jacobson C.;
RT "Lysine residue 240 of human serum albumin is involved in high-affinity binding of bilirubin.";
RL Biochem. J. 171:453-459 (1978).
[16] VARIANTS NAG-2 AND NAG-3.
RA MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197 (1987).
[17] VARIANTS NAG-2 AND NAG-3.
RA MEDLINE=89068523; PubMed=3479777;
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M., Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from Amerindian and Japanese populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005 (1987).
[18] VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RA MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M., Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese allcoalbumins.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096 (1989).
[19] VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RA MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501 (1990).
[20] DESCRIPTION OF VARIANT REDHILL.
RA MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of human serum albumin whose precursor has an aberrant signal peptidase cleavage site.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30 (1990).
[21] VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RA MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J., Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in Italy.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725 (1990).
[22] VARIANT VENEZIA.
RA MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M., Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two carboxyl-terminal variants of human serum albumin.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963 (1991).
[23] VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23; KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RA MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E., Matsuda Y.-I., Anaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857 (1991).

[24] VARIANT CASEBROOK ASN-518.
RA MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum albumin: albumin Casebrook (494 Asp-->Asn).";
RL Biochim. Biophys. Acta 1097:49-54 (1991).
[25] VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RA MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H., Rochu D., Porta F.;
RT "Two allcoalbumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 57; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKKE 12
DB 290 ENQDSISSKKE 301

RESULT 3
ALBU CANFA
ID ALBU CANFA STANDARD; PRT; 608 AA.
AC P49822; 077705; 09TS24;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).
GN ALB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285 (2000).
RN [3]
RP SEQUENCE OF 25-48.
RA MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W., Sarkar B.;
RT "Isolation, amino acid sequence and copper(II)-binding properties of peptide (1-24) of dog serum albumin.";
RL J. Biol. Chem. 249:5872-5877 (1974).
RN [4]
RP SEQUENCE OF 25-38.
RC TISSUE=Heart;
RA MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of dog heart proteins.";
RL Electrophoresis 18:2795-2802 (1997).
RN [5]
RP SEQUENCE OF 215-478 FROM N.A.
RC TISSUE=Salivary gland;
RA MEDLINE=94201492; PubMed=7512102;
RA Spitzauer S., Schweizer C., Sperr W.R., Pandjaitan B., Valent P., Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
RT "Molecular characterization of dog albumin as a cross-reactive allergen.";

RESULT 4	ALBU_FELCA	STANDARD;	PRT;	608 AA.
ID	ALBU_FELCA			
AC	P49064;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Serum albumin precursor (Allergen Fel d 2).			
GN	Ala.			
OS	Felis silvestris catus (Cat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.			
OX	NCBI_TaxID=9685;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96194824; PubMed=8647469;			
RA	Hilger C., Grigioni F., Kohlen M., Hentges F.;			
RT	"Sequence of the gene encoding cat (Felis domesticus) serum albumin.";			
RL	Gene 1691295-296(1996).			
CC	-!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- TISSUE SPECIFICITY: Plasma.			
CC	-!- ALLERGEN: Causes an allergic reaction in human.			
CC	-!- SIMILARITY: Belongs to the ALB/AFP/VDS family.			
CC	-!- SIMILARITY: Contains 3 albumin domains.			
CC	-!- SIMILARITY: Contains 3 albumin domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; X84842; CAA59279.1; --			
DR	PIR; JC4660; S57632.			
DR	HSP; P02768; 1B7B.			
DR	InterPro; IPR000264; Serum_albumin.			
DR	Pfam; PF00273; transport_prot; 3.			
DR	PRINTS; PR00802; SERUMALBUMIN.			
DR	ProDom; PD002486; Serum_albumin; 1.			
DR	SMART; SM00103; ALBUMIN; 3.			
DR	Metal-binding; PS00212; ALBUMIN; 3.			
KW	Metals-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.			
FT	SIGNAL 1 18 BY SIMILARITY.			
FT	PROPEP 19 24 BY SIMILARITY.			
FT	CHAIN 25 608 SERUM ALBUMIN.			
FT	DOMAIN 25 205 ALBUMIN 1.			
FT	DOMAIN 212 397 ALBUMIN 2.			
FT	DOMAIN 404 595 ALBUMIN 3.			
FT	METAL 27 27 COPPER.			
FT	DISULFID 77 86 BY SIMILARITY.			
FT	DISULFID 99 115 BY SIMILARITY.			
FT	DISULFID 114 125 BY SIMILARITY.			
FT	DISULFID 148 193 BY SIMILARITY.			
FT	DISULFID 192 201 BY SIMILARITY.			
FT	DISULFID 224 270 BY SIMILARITY.			
FT	DISULFID 269 277 BY SIMILARITY.			
FT	DISULFID 289 303 BY SIMILARITY.			
FT	DISULFID 302 313 BY SIMILARITY.			
FT	DISULFID 340 385 BY SIMILARITY.			
FT	DISULFID 384 393 BY SIMILARITY.			
FT	DISULFID 416 462 BY SIMILARITY.			
FT	DISULFID 461 472 BY SIMILARITY.			
FT	DISULFID 485 501 BY SIMILARITY.			
FT	DISULFID 500 511 BY SIMILARITY.			
FT	DISULFID 538 583 BY SIMILARITY.			
FT	DISULFID 582 591 BY SIMILARITY.			


```
Query Match          94.7%; Score 54; DB 1; Length 608;
Best Local Similarity 91.7%; Pred. No. 0.028;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDISISKLKE 12
DB 290 ENQDISISKLKE 301

RESULT 5
ALBU_FIG STANDARD; PRT; 605 AA.
AC P08835; Q29018;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S., Weinstock J.;
RT "Nucleotide sequence of porcine liver albumin.";
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12422; CAA30970.1; -.
DR EMBL; M35787; AAA30988.1; -.
DR PIR; S01382; ABPGS.
DR HSP; P02768; 1E7H.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_P1ot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON_TER 1
FT SIGNAL 16 BY SIMILARITY.
FT PROPEP 17 22 BY SIMILARITY.
FT CHAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.
FT DOMAIN 401 592 ALBUMIN 3.
FT METAL 31 31 COPPER (BY SIMILARITY).
FT METAL 75 84 COPPER (BY SIMILARITY).
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.
```

```
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1F4FF CRC64;

Query Match          89.5%; Score 51; DB 1; Length 605;
Best Local Similarity 83.3%; Pred. No. 0.1;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDISISKLKE 12
DB 287 ENQDISISKLKE 298

RESULT 6
ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Holowachuk E.W., Stoltenberg J.K., Reed R.G., Peters T. Jr.;
RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RC TISSUE=Liver;
RA Barry T., Power S., Cannon F.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Halger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T., Huang M.C.;
RT "The complete cDNA sequence of bovine serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-32.
RX MEDLINE=80024278; PubMed=488109;
RA McGillivray R.T.A., Chung D.W., Davie E.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
RT terminal sequence of prealbumin.";
RL Eur. J. Biochem. 98:477-485(1979).
RN [6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA Brown J.R.;
RT "Structure of bovine serum albumin.";
RL Fed. Proc. 34:591-591(1975).
RN [7]
RP REVISIONS TO 190-195.
RA Brown J.R.;
RL Submitted (APR-1975) to the PIR data bank.
RN [8]
RP SEQUENCE OF 402-433.
```

RX MEDLINE=82023364; PubMed=7283978;
RA Reed R.G., Putnam F.W., Peters T. Jr.;
RT "Sequence of residues 400-403 of bovine serum albumin.";
RL Biochem. J. 191:867-868(1980).
RN [9]
RP SEQUENCE OF 19-28.
RX MEDLINE=77134075; PubMed=843354;
RA Patterson J.E., Geller D.M.;
RT "Bovine microsomal albumin: amino terminal sequence of bovine proalbumin.";
RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
RN [10]
RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.
RX MEDLINE=91083649; PubMed=2260975;
RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;
RT "Rapid confirmation and revision of the primary structure of bovine serum albumin by ESMS and Frit-FAB LC/MS.";
RL Biochem. Biophys. Res. Commun. 173:639-646(1990).
RN [11]
RP SEQUENCE OF 25-41.
RX MEDLINE=88267456; PubMed=3389500;
RA Hsieh J.C., Lin F.P., Tam M.F.;
RT "Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel: a preparative method for isolating proteins for N-terminal microsequencing.";
RL Anal. Biochem. 170:1-8(1988).
RN [12]
RP SEQUENCE OF 437-451.
RA Vilbois F.;
RL Submitted (AUG-1998) to Swiss-Prot.
RN [13]
RP DISULFIDE BONDS.
RA Brown J.R.;
RL "Structure of serum albumin: disulfide bridges.";
RL Fed. Proc. 33:1389-1389(1974).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR ENBL; M73993; AAA51411.1; -;
DR ENBL; X58989; CAA41735.1; -;
DR ENBL; Y17769; CAA76847.1; -;
DR ENBL; AF542068; AAN17824.1; -;
DR HSP; P02768; 1E7B.
DR InterPro: I2R00264; Serum albumin.
DR Pfam: PF00273; Transport_prot; 3.
DR PRINTS; P00802; SERUMALBUMIN.
DR ProDom; P002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
KW Polymorphism.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).

FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
FT VARIANT 214 214
FT CONFLICT 302 302
FT CONFLICT 304 305
FT CONFLICT 324 324
FT CONFLICT 394 395
FT CONFLICT 437 437
FT CONFLICT 493 494
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE78585D4 CRC64;
Query Match 89.5%; Score 51; DB 1; Length 607;
Best Local Similarity 83.3%; Pred. No. 0.1;
Matches 10; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 ENQDSISSKLE 12
Db 289 DNQDTISSKLE 300

RESULT 7
ID MANA HUMAN STANDARD; PRT; 422 AA.
AC P34949;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mannose-6-phosphate isomerase (EC 5.3.1.8) (Phosphomannose isomerase) (MPI) (Phosphohexomutase).
GN MPI OR PMI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Placenta, and Testis;
RX MEDLINE=94139717; PubMed=8307007;
RA Proudfoot A.E.I., Turcatti G., Wells T.N.C., Payton M.A., Smith D.J.;
RT "Purification, cDNA cloning and heterologous expression of human phosphomannose isomerase.";
RL Eur. J. Biochem. 219:415-423(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20438347; PubMed=10980531;
RA Scholien E., Dorland L., de Koning T.J., Van Diggelen O.P., Patterson M., Huijman J.G.M., Marquardt T., Babovic-Vuksanovic D., Pateron M., Intiaz F., Winchester B., Adamowicz M., Pronicka E., Freeze H., Matthijs G.;
RT "Genomic organization of the human phosphomannose isomerase (MPI) gene and mutation analysis in patients with congenital disorders of glycosylation type Ib (CDG-Ib).";
RL Hum. Mutat. 16:247-252(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Narusins K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzly D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grunwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
[4]
RP VARIANTS CDG-IB LEU-101 AND THR-137.
RX MEDLINE=98254476; PubMed=9585601;
RA Jaeken J., Matthijs G., Saudubray J.-M., Dionisi-Vici C., Bertini E.,
RA de Lonlay P., Henri H., Carchon H., Schollen E., Van Schaftingen E.,
RT "Phosphomannose isomerase deficiency: a carbohydrate-deficient
RT glycoprotein syndrome with hepatic-intestinal presentation.";
RL Am. J. Hum. Genet. 62:1535-1539 (1998).
[5]
RP VARIANT CDG-IB GLN-218.
RX MEDLINE=98192611; PubMed=9525984;
RA Niehus R., Hasilik M., Alton G., Koerner C., Schliebe-Sukumar M.,
RA Koch H.G., Zimmer K.-P., Wu R., Harms E., Reiter K., von Figura K.,
RA Freeze H.H., Harms H.K., Marquardt T.,
RT "Carbohydrate-deficient glycoprotein syndrome type Ib: phosphomannose
RT isomerase deficiency and mannose therapy.";
RL J. Clin. Invest. 101:1414-1420 (1998).
[6]
RP VARIANTS CDG-IB THR-139 AND GLN-218.
RX MEDLINE=21249093; PubMed=11350186;
RA Westphal V., Kjaergaard S., Davis J.A., Peterson S.M., Skovby F.,
RA Freeze H.H.,
RT "Genetic and metabolic analysis of the first adult with congenital
RT disorder of glycosylation type Ib: long-term outcome and effects of
RT mannose supplementation.";
RL Mol. Genet. Metab. 73:77-85 (2001).
CC -!- FUNCTION: Involved in the synthesis of the GDP-mannose and
CC dolichol-phosphate-mannose required for a number of critical
CC mannosiyl transfer reactions.
CC -!- CATALYTIC ACTIVITY: D-mannose 6-phosphate = D-fructose 6-
CC phosphate.
CC -!- COFACTOR: Zinc (By similarity).
CC -!- PATHWAY: Glycosylation; early steps of mannosylation.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- TISSUE SPECIFICITY: Expressed in all tissues, but more abundant in
CC heart, brain and skeletal muscle.
CC -!- DISEASE: Defects in MPI are the cause of congenital disorder of
CC glycosylation type Ib (CDG-Ib) (MIM:602579); also known as
CC carbohydrate-deficient glycoprotein syndrome type IB (CDGS-IB).
CC CDG-IB is clinically characterized by protein-losing enteropathy,
CC a gastrointestinal disorder. Biochemically it is characterized by
CC hypoglycosylation of serum glycoproteins.
CC -!- SIMILARITY: Belongs to the mannose-6-phosphate isomerase family 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X76057; CAA53657.1; -.
DR EMBL; AF227218; AAF37697.1; -.

DR EMBL; AF227216; AAF37697.1; JOINED.
DR EMBL; AF227217; AAF37697.1; JOINED.
DR EMBL; BC046357; AAH46357.1; -.
DR F1R; S41122; S41122.
DR HSSP; P34948; LPM1.
DR Genew; HGNC:7216; MPI.
DR MIM; 154550; -.
DR GO; GO:0004476; F:mannose-6-phosphate isomerase activity; TAS.
DR InterPro; IPR001250; Man6p_isomerase1.
DR Pfam; PF01238; PMI type1; I.
DR PRINTS; PR00714; MAN6PISMRASE.
DR PRODOM; PD004391; Man6p_isomerase1; 1.
DR TIGRFAMs; TIGR00218; man6a; 1.
DR PROSITE; PS00965; PMI_1; 1.
DR PROSITE; PS00966; PMI_1.2; 1.
KW Isomerase; Zinc; Disease mutation.
FT INIT MET 0 PROBABLE.
FT METAL 109 109 ZINC (BY SIMILARITY).
FT METAL 111 111 ZINC (BY SIMILARITY).
FT METAL 136 136 ZINC (BY SIMILARITY).
FT METAL 275 275 ZINC (BY SIMILARITY).
FT VARIANT 101 101 S -> L (in CDG-IB).
FT VARIANT 137 137 /FTID=VAR_012338.
FT VARIANT 139 139 M -> T (in CDG-IB).
FT VARIANT 139 139 I -> T (in CDG-IB).
FT VARIANT 218 218 /FTID=VAR_012345.
FT VARIANT 218 218 R -> Q (in CDG-IB).
FT VARIANT 218 218 /FTID=VAR_012340.
SQ SEQUENCE 422 AA; 46524 MW; A450ABBF53722605 CRC64;
Query Match 78.9%; Score 45; DB 1; Length 422;
Best Local Similarity 66.7%; Pred. No. 0.88;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ENQDSISSKLE 12
Db 76 ENQDSLGSVKVD 87
|||||:|:|:
RESULT 8
AFAM_HUMAN STANDARD; PRT; 599 AA.
ID AFAM_HUMAN
AC P43652;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE AFamin precursor (Alpha-albumin) (Alpha-Alb).
GN AFM OR ALBA OR ALB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC Tissue=Liver;
RX MEDLINE=9429534; PubMed=7517938;
RA Lichenstein H.S., Lyons D.E., Wurfel M.M., Johnson D.A.,
RA McGinley M.D., Leidl J.C., Trollinger D.B., Mayer J.P.,
RA Wright S.D., Zukowski M.M.,
RT "AfaMin is a new member of the albumin, alpha-fetoprotein, and
RT vitamin D-binding protein gene family.";
RL J. Biol. Chem. 269:18149-18154 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96240683; PubMed=8648639;
RA Nishio H., Heiskanen M., Palotie A., Belanger L., Dugaiczak A.,
RT "Random arrangement of the human serum albumin multigene family in
RT the sub-centromeric region of 4q: evolution and chromosomal direction
RT of transcription.";
RL J. Mol. Biol. 259:113-119 (1996).
RN [3]

```
RP SEQUENCE OF 1-69; 105-207 AND 560-599 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95180738; PubMed=7875606;
RA Allard D., Gilbert S., Lamontagne A., Hamel D., Belanger L.;
RT "Identification of rat alpha-albumin and cDNA cloning of its human
ortholog";
RL Gene 153:287-288(1995).
CC -!- FUNCTION: Possible role in the transport of yet unknown ligand.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: N-glycosylated.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32140; AAA21612.1; -
DR EMBL; U51243; AAC50720.1; -
DR EMBL; L35486; AAA68197.1; -
DR EMBL; L35497; AAA68198.1; -
DR EMBL; L35498; AAA68199.1; -
DR PIR; A54906; A54906
DR PIR; I39424; I39424
DR PIR; I39424; I39424
DR PIR; I39426; I39426
DR HSP; P02768; IE7B.
DR Genew; HGNC:316; AFM.
DR MIM; 104145; -
DR GO; GO:0005576; C:extracellular; TAS.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Transport; Repeat; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 599
FT DOMAIN 22 205
FT DOMAIN 212 397
FT DOMAIN 404 593
FT DISULFID 77 86
FT DISULFID 113 124
FT DISULFID 148 193
FT DISULFID 197 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 470
FT DISULFID 483 499
FT DISULFID 498 509
FT DISULFID 536 581
FT DISULFID 580 589
FT CARBOHYD 33 33
FT CARBOHYD 109 109
FT CARBOHYD 402 402
FT CARBOHYD 488 488
SQ SEQUENCE 599 AA; 69065 MW; D594E75E20D308AB CRC64;

Query Match
Best Local Similarity 90.0%; Pred. No. 1,9; Length 599;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 QDSISSKLKE 12
```

```
DB DB 292 QDSISSKLKE 301
RESULT 9
ALBU SHEEP
ID ALBU SHEEP STANDARD; PRT; 607 AA.
AC F14639;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90095888; PubMed=2602160;
RA Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
RL Nucleic Acids Res. 17:10495-10495(1989).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17055; CAA34903.1; -
DR PIR; S06936; ABSHS.
DR HSP; P02768; IE7B.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT CHAIN 19 24
FT CHAIN 25 607
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
```

```

FT DISULFID 499 510 BY SIMILARITY.
FT DISULFID 537 582 BY SIMILARITY.
FT DISULFID 581 590 BY SIMILARITY.
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 77.2%; Score 44; DB 1; Length 607;
Best Local Similarity 66.7%; Pred. No. 2;
Matches 8; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKLXE 12
: : : : :
Db 289 DHQDALSCLKX 300

RESULT 10
ALBU HORSE
ID ALBU HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]_TaxID=9796;
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)
at 0.27-nm resolution."
RL Eur. J. Biochem. 215:205-212(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IGF.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74045; CAA52194.1; -
CC PIR; S34053; ABHOS.
CC -----
CC DR HSP; P02768; 1E7B.
CC DR InterPro; IPR000284; Serum_albumin.
CC DR Pfam; PF00273; transport_prot; 3.
CC DR PRINTS; PR00802; SERUMALBUMIN.
CC DR ProDom; PD002486; Serum_albumin; 1.
CC DR SMART; SM00103; ALBUMIN; 3.
CC DR PROSITE; PS00212; ALBUMIN; 3.
CC KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
CC FT SIGNAL 1 18
CC FT PROPEP 19 24
CC FT CHAIN 25 607 BY SIMILARITY.
CC FT DOMAIN 25 204 SERUM ALBUMIN.
CC FT DOMAIN 211 396 ALBUMIN 1.
CC FT DOMAIN 403 594 ALBUMIN 2.
CC FT METAL 27 27 COPPER (BY SIMILARITY).
CC FT DISULFID 77 86
CC FT DISULFID 99 115
CC FT DISULFID 114 125

```

```

FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;

Query Match 75.4%; Score 43; DB 1; Length 607;
Best Local Similarity 81.8%; Pred. No. 3;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKLK 11
: : : : :
Db 289 EHQDSISGKLK 299

RESULT 11
ALBU MERUN
ID ALBU MERUN STANDARD; PRT; 609 AA.
AC Q35030;
DT 13-JUL-1998 (Rel. 36, Created)
DT 13-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OX NCBI_TaxID=10047;
RN [1]_TaxID=10047;
RP SEQUENCE FROM N.A.
RC STRAIN=MGS IDR; TISSUE=Liver;
RX MEDLINE=98116663; PubMed=9455485;
RA Yoshida K., Seto-Oshima A., Sinochura H.;
RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
synthesis in the Mongolian gerbil, Meriones unguiculatus."
RL DNA Res. 4:351-354(1997).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB006197; BAA21765.1; -
CC PIR; JCS838; JCS838.
CC DR HSP; P02768; 1E7B.
CC DR InterPro; IPR000264; Serum_albumin.
CC DR Pfam; PF00273; transport_prot; 3.
CC DR PRINTS; PR00802; SERUMALBUMIN.
CC DR ProDom; PD002486; Serum_albumin; 1.
CC DR SMART; SM00103; ALBUMIN; 3.
CC DR PROSITE; PS00212; ALBUMIN; 3.

```

```
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 18 BY SIMILARITY.
FT PROPEP 19 24
FT CHAIN 25 609
FT SERUM ALBUMIN.
FT DOMAIN 25 206
FT ALBUMIN 1.
FT DOMAIN 213 398
FT ALBUMIN 2.
FT DOMAIN 405 596
FT ALBUMIN 3.
FT METAL 28 28
FT COPPER.
FT DISULFID 78 87
FT BY SIMILARITY.
FT DISULFID 100 116
FT BY SIMILARITY.
FT DISULFID 115 126
FT BY SIMILARITY.
FT DISULFID 149 194
FT BY SIMILARITY.
FT DISULFID 193 202
FT BY SIMILARITY.
FT DISULFID 225 271
FT BY SIMILARITY.
FT DISULFID 270 278
FT BY SIMILARITY.
FT DISULFID 290 304
FT BY SIMILARITY.
FT DISULFID 303 314
FT BY SIMILARITY.
FT DISULFID 341 386
FT BY SIMILARITY.
FT DISULFID 385 394
FT BY SIMILARITY.
FT DISULFID 417 463
FT BY SIMILARITY.
FT DISULFID 462 473
FT BY SIMILARITY.
FT DISULFID 486 502
FT BY SIMILARITY.
FT DISULFID 501 512
FT BY SIMILARITY.
FT DISULFID 539 584
FT BY SIMILARITY.
FT DISULFID 583 592
FT BY SIMILARITY.
SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F57E1A48 CRC64;

Query Match 70.2%; Score 40; DB 1; Length 609;
Best Local Similarity 81.8%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKLK 11
Db 291 ENQASISSKLQ 301

RESULT 12
TEGU_HSVSA STANDARD; PRT; 2469 AA.
AC Q01056;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Probable large tegument protein.
GN 64 OR ERF2.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gamaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10393;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92233688; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome."
RL J. Virol. 66:5047-5059 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92230228; PubMed=1314457;
RA Nicholas J., Cameron K.R., Coleman H., Newman C., Honess R.W.;
RA "Analysis of nucleotide sequence of the rightmost 43 kbp of
RT herpesvirus saimiri (HVS) L-DNA: general conservation of genetic
RT organization between HVS and Epstein-Barr virus."
RL Virology 188:296-310 (1992).
CC -!- FUNCTION: Tegument protein.
CC BHV-1 24, EBV BPLF1, HVS-1 64, VZV 22, AND HCMV UL48.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X76456; CAA53994.1; -
PIR; A53195; A53195.
HSP; P02768; I37B.
InterPro; IPR000264; Serum albumin.
Pfam; PF00273; transport prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
PRODOM; PD002486; Serum albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 1.
Transprot; Repeat; Glycoprotein; Signal.
FT CHAIN 21 21 BY SIMILARITY.
FT SIGNAL 22 608
FT DOMAIN 22 205
FT DOMAIN 212 397
FT DOMAIN 404 593
FT DISULFID 77 86
FT DISULFID 99 114
FT DISULFID 113 124
FT DISULFID 148 193
FT DISULFID 224 270
FT DISULFID 269 277

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 18 BY SIMILARITY.
FT PROPEP 19 24
FT CHAIN 25 609
FT SERUM ALBUMIN.
FT DOMAIN 25 206
FT ALBUMIN 1.
FT DOMAIN 213 398
FT ALBUMIN 2.
FT DOMAIN 405 596
FT ALBUMIN 3.
FT METAL 28 28
FT COPPER.
FT DISULFID 78 87
FT BY SIMILARITY.
FT DISULFID 100 116
FT BY SIMILARITY.
FT DISULFID 115 126
FT BY SIMILARITY.
FT DISULFID 149 194
FT BY SIMILARITY.
FT DISULFID 193 202
FT BY SIMILARITY.
FT DISULFID 225 271
FT BY SIMILARITY.
FT DISULFID 270 278
FT BY SIMILARITY.
FT DISULFID 290 304
FT BY SIMILARITY.
FT DISULFID 303 314
FT BY SIMILARITY.
FT DISULFID 341 386
FT BY SIMILARITY.
FT DISULFID 385 394
FT BY SIMILARITY.
FT DISULFID 417 463
FT BY SIMILARITY.
FT DISULFID 462 473
FT BY SIMILARITY.
FT DISULFID 486 502
FT BY SIMILARITY.
FT DISULFID 501 512
FT BY SIMILARITY.
FT DISULFID 539 584
FT BY SIMILARITY.
FT DISULFID 583 592
FT BY SIMILARITY.
SQ SEQUENCE 609 AA; 68940 MW; 9CA5F97F57E1A48 CRC64;

Query Match 70.2%; Score 40; DB 1; Length 2469;
Best Local Similarity 81.8%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENQDSISSKLK 12
Db 373 EDDDNVTSKLKE 384

RESULT 13
AFAM_RAT STANDARD; PRT; 608 AA.
ID AFAM_RAT
AC P36953;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Afamin precursor (Alpha-albumin) (Alpha-Alb).
GN AFM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley;
RX MEDLINE=94164881; PubMed=7509788;
RA Belanger L., Roy S., Allard D.;
RT "New albumin gene 3' adjacent to the alpha 1-fetoprotein locus.";
RL J. Biol. Chem. 269:5481-5484 (1994).
CC -!- FUNCTION: Possible role in the transport of yet unknown ligand.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X76456; CAA53994.1; -
PIR; A53195; A53195.
HSP; P02768; I37B.
InterPro; IPR000264; Serum albumin.
Pfam; PF00273; transport prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
PRODOM; PD002486; Serum albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 1.
Transprot; Repeat; Glycoprotein; Signal.
FT CHAIN 21 21 BY SIMILARITY.
FT SIGNAL 22 608
FT DOMAIN 22 205
FT DOMAIN 212 397
FT DOMAIN 404 593
FT DISULFID 77 86
FT DISULFID 99 114
FT DISULFID 113 124
FT DISULFID 148 193
FT DISULFID 224 270
FT DISULFID 269 277
```

FT DISULFID 289 303 BY SIMILARITY. ALBUMIN 2.
FT DISULFID 302 313 BY SIMILARITY. ALBUMIN 3.
FT DISULFID 340 385 BY SIMILARITY. COPPER.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 470 BY SIMILARITY.
FT DISULFID 483 499 BY SIMILARITY.
FT DISULFID 498 509 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 608 AA; 693335 MW; F33151A6B8A07F6 CRC64;
Query Match 68.4%; Score 39; DB 1; Length 608;
Best Local Similarity 89.9%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 3 QDSISSKXK 11
Db 292 QDSISSKXK 300
RESULT 14
ALBU_RABIT
ID ALBU_RABIT STANDARD; PRT; 608 AA.
AC P49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-New Zealand white; TISSUE=Liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U18344; AAB58347.1; .
DR HSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT

FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 495 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68914 MW; CF5E92647AAFE9A2 CRC64;
Query Match 68.4%; Score 39; DB 1; Length 608;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ENQDSISSKXK 12
Db 290 ENQETISSHLK 301
RESULT 15
AFAM_MOUSE
ID AFAM_MOUSE STANDARD; PRT; 611 AA.
AC O89020;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Afamin precursor (Alpha-albumin) (Alpha-Alb).
GN AFM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Diaphragm;
RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Possible role in the transport of yet unknown ligand.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AJ011080; CAA09471.1; .
DR HSP; P02768; 1E7B.
DR MG; MG1:2429409; Afm.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
KW Metal-binding; Repeat; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT

FT CHAIN	22	611	AFAMIN.
FT DOMAIN	22	205	ALBUMIN 1.
FT DOMAIN	212	397	ALBUMIN 2.
FT DOMAIN	404	593	ALBUMIN 3.
FT DISULFID	77	86	BY SIMILARITY.
FT DISULFID	99	114	BY SIMILARITY.
FT DISULFID	113	124	BY SIMILARITY.
FT DISULFID	148	193	BY SIMILARITY.
FT DISULFID	224	270	BY SIMILARITY.
FT DISULFID	269	277	BY SIMILARITY.
FT DISULFID	289	303	BY SIMILARITY.
FT DISULFID	302	313	BY SIMILARITY.
FT DISULFID	340	385	BY SIMILARITY.
FT DISULFID	384	393	BY SIMILARITY.
FT DISULFID	416	462	BY SIMILARITY.
FT DISULFID	461	470	BY SIMILARITY.
FT DISULFID	483	499	BY SIMILARITY.
FT DISULFID	498	509	BY SIMILARITY.
FT DISULFID	580	589	BY SIMILARITY.
FT CARBOHYD	33	33	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	109	109	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	153	153	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	402	402	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	488	488	N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE	611 AA;	69635 MW;	39E46B6E723F89C8 CRC64;

Query Match 68.4%; Score 39; DB 1; Length 611;
 Best Local Similarity 88.9%; Pred. No. 16;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	3 QDSISSKLK 11
Db	292 QDSISSKIX 300

Search completed: April 19, 2004, 11:52:51
 Job time : 1.87535 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 4.63158 Seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_266_277

Perfect score: 57
Sequence: 1 ENQDS-SSKLKE 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_xvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	100.0	417	4 Q85YGO	Q85YGO homo sapien
2	54	94.7	584	6 Q7YSG3	Q7YSG3 felis silve
3	45	78.9	267	4 Q8NHZ6	Q8NHZ6 homo sapien
4	45	78.9	362	4 Q95AB0	Q95AB0 homo sapien
5	45	78.9	423	6 Q8HX2	Q8HX2 macaca fasc
6	41	71.9	74	2 Q7WVJ9	Q7WVJ9 uncultured
7	41	71.9	74	2 Q7WVJ7	Q7WVJ7 uncultured
8	41	71.9	74	2 Q7WVJ6	Q7WVJ6 uncultured
9	41	71.9	74	2 Q7WVH6	Q7WVH6 uncultured
10	41	71.9	74	2 Q7WVH4	Q7WVH4 uncultured
11	41	71.9	74	2 Q7WVH3	Q7WVH3 uncultured
12	41	71.9	74	2 Q7WVG9	Q7WVG9 uncultured
13	41	71.9	74	2 Q7WRK7	Q7WRK7 uncultured
14	41	71.9	74	2 Q7WRK6	Q7WRK6 uncultured
15	41	71.9	74	2 Q7WRK5	Q7WRK5 uncultured
16	40	70.2	74	2 Q7WVK4	Q7WVK4 uncultured

17	40	70.2	74	2	Q7WVK1	Q7WVK1 uncultured
18	40	70.2	74	2	Q7WVJ8	Q7WVJ8 uncultured
19	40	70.2	74	2	Q7WVJ6	Q7WVJ6 uncultured
20	40	70.2	74	2	Q7WVJ5	Q7WVJ5 uncultured
21	40	70.2	74	2	Q7WVJ4	Q7WVJ4 uncultured
22	40	70.2	74	2	Q7WVJ3	Q7WVJ3 uncultured
23	40	70.2	74	2	Q7WVJ1	Q7WVJ1 uncultured
24	40	70.2	74	2	Q7WVJ7	Q7WVJ7 uncultured
25	40	70.2	74	2	Q7WVJ5	Q7WVJ5 uncultured
26	40	70.2	74	2	Q7WVJ3	Q7WVJ3 uncultured
27	40	70.2	74	2	Q7WVJ2	Q7WVJ2 uncultured
28	40	70.2	74	2	Q7WVH8	Q7WVH8 uncultured
29	40	70.2	74	2	Q7WVH7	Q7WVH7 uncultured
30	40	70.2	74	2	Q7WVH5	Q7WVH5 uncultured
31	40	70.2	74	2	Q7WVH2	Q7WVH2 uncultured
32	40	70.2	74	2	Q7WVH0	Q7WVH0 uncultured
33	40	70.2	74	2	Q7WRK8	Q7WRK8 uncultured
34	40	70.2	74	2	Q7WRK4	Q7WRK4 uncultured
35	40	70.2	74	2	Q7WRK3	Q7WRK3 uncultured
36	40	70.2	608	5	Q95VB7	Q95VB7 schistosoma
37	40	70.2	2477	12	Q80BL6	Q80BL6 saimirine
38	39	68.4	74	2	Q7WVK8	Q7WVK8 uncultured
39	39	68.4	74	2	Q7WVK5	Q7WVK5 uncultured
40	39	68.4	74	2	Q7WVK3	Q7WVK3 uncultured
41	39	68.4	74	2	Q7WVK2	Q7WVK2 uncultured
42	39	68.4	74	2	Q7WVI9	Q7WVI9 uncultured
43	39	68.4	74	2	Q7WVI8	Q7WVI8 uncultured
44	39	68.4	74	2	Q7WVH9	Q7WVH9 uncultured
45	39	68.4	74	2	Q7WVG6	Q7WVG6 uncultured

ALIGNMENTS

RESULT 1
Q85YGO PRELIMINARY; PRT; 417 AA.
ID Q85YGO
AC Q85YGO; (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Srausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1;
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; P:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; S:Serum albumin.
DR Pfam; PF00273; t:transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; S:Serum albumin; 1.
DR SMART; SMC0103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47360 MW; 16E764833EEF4E8D CRC64;

Query Match 100.0%; Score 57; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 0.056; 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Indels 0;
QY 1 ENQDSISSKLKE 12
Db 98 ENQDSISSKLKE 109

RESULT 2

Q7YSG3 Q7YSG3 PRELIMINARY; PRT; 584 AA.
AC Q7YSG3, 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Albumin (Fragment).
ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA Rumpold H., Valenta R., Spitzauer S.,
RT "Escherichia coli expression and purification of recombinant cat
albumin: IgE recognition, induction of basophil activation and
lymphoproliferative responses in atopic patients.";
RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ487677; CAD32275.1; -.
FT NON TER 1
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;
Query Match 94.7%; Score 54; DB 6; Length 584;
Best Local Similarity 91.7%; Pred. No. 0.27; Mismatches 1; Indels 0; Gaps 0;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 ENQDSISXKKE 12
Db 266 ENQDSISXKKE 277
RESULT 3
Q8NHZ6 PRELIMINARY; PRT; 267 AA.
AC Q8NHZ6;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Mannose phosphate isomerase isoform (EC 5.3.1.8) (Mannose-6-phosphate
isomerase) (Phosphomannose isomerase) (PMI)
DE (Phosphohexomutase).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Guo J.H., Yu L.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-MANNOSE 6-PHOSPHATE = D-FRUCTOSE 6-
PHOSPHATE.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MANNOSE-6-PHOSPHATE ISOMERASE FAMILY 1.
DR EMBL; AF504648; AM28199.1; -.
DR GO; GO:0016953; F:isomerase activity; IEA.
DR GO; GO:0004476; F:mannose-6-phosphate isomerase activity; IEA.
DR GO; GO:0008270; P:zinc ion binding; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001250; Man6p_isomerase.
DR Pfam; PF01238; PMI_type1; 1.
DR PRINTS; PR00714; MAN6PISMRASE.
DR ProDom; PD004391; Man6p_isomerase; 1.
DR TIGRFAMs; TIGR00218; manA; 1.
DR PROSITE; PS00965; PMI_I_1; 1.
DR PROSITE; PS00966; PMI_I_2; 1.
KW Isomerase; Zinc.
SQ SEQUENCE 267 AA; 29699 MW; 322DCD06F3F92FE CRC64;
Query Match 78.9%; Score 45; DB 4; Length 267;
Best Local Similarity 66.7%; Pred. No. 5.3; Mismatches 3; Indels 0; Gaps 0;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 ENQDSISXKKE 12
Db 77 ENQDSISXKKE 88
RESULT 5
Q8HXX2 PRELIMINARY; PRT; 423 AA.
AC Q8HXX2;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Mannose-6-phosphate isomerase.
DE MPI.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cerebellum cortex;
RA Kusuda J., Osada N., Hida M., Sugano S., Hashimoto K.;
RT "Isolation and characterization of cDNA for macaque neurological
disease genes.";
RT

QY 1 ENQDSISXKKE 12
Db 57 ENQDSISXKKE 68
RESULT 4
Q96AB0 PRELIMINARY; PRT; 362 AA.
AC Q96AB0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to mannose phosphate isomerase (EC 5.3.1.8) (Mannose-6-
phosphate isomerase) (Phosphomannose isomerase) (PMI)
DE (Phosphohexomutase).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Astrocytoma;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-MANNOSE 6-PHOSPHATE = D-FRUCTOSE 6-
PHOSPHATE.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE MANNOSE-6-PHOSPHATE ISOMERASE FAMILY 1.
DR EMBL; BC017351; AAH17351.1; -.
DR GO; GO:0016953; F:isomerase activity; IEA.
DR GO; GO:0004476; F:mannose-6-phosphate isomerase activity; IEA.
DR GO; GO:0008270; P:zinc ion binding; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001250; Man6p_isomerase.
DR Pfam; PF01238; PMI_type1; 1.
DR PRINTS; PR00714; MAN6PISMRASE.
DR ProDom; PD004391; Man6p_isomerase; 1.
DR TIGRFAMs; TIGR00218; manA; 1.
DR PROSITE; PS00965; PMI_I_1; 1.
DR PROSITE; PS00966; PMI_I_2; 1.
KW Isomerase; Zinc.
SQ SEQUENCE 362 AA; 39834 MW; 58CA9B39BF20C459 CRC64;
Query Match 78.9%; Score 45; DB 4; Length 362;
Best Local Similarity 66.7%; Pred. No. 7.1; Mismatches 3; Indels 0; Gaps 0;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 ENQDSISXKKE 12
Db 77 ENQDSISXKKE 88
RESULT 5
Q8HXX2 PRELIMINARY; PRT; 423 AA.
AC Q8HXX2;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Mannose-6-phosphate isomerase.
DE MPI.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cerebellum cortex;
RA Kusuda J., Osada N., Hida M., Sugano S., Hashimoto K.;
RT "Isolation and characterization of cDNA for macaque neurological
disease genes.";
RT

Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
EMBL: AB083318; BAC20597.1; -
GO: GO:0015853; P:isomerase activity; IEA.
GO: GO:0004470; P:mannose-6-phosphate isomerase activity; IEA.
GO: GO:0008270; P:zinc ion binding; IEA.
GO: GO:0005975; P:carbohydrate metabolism; IEA.
InterPro: IPR001250; Man6p_isomerase.
Pfam: PF01238; PMI type1; 1.
PRINTS: PR00714; MANGPISNRASE.
ProDom: PD004391; Man6p_isomerase1; 1.
TIGRFAMs: TIGR00218; man6a; 1.
PROSITE: PS00965; PMI_1; 1.
PROSITE: PS00966; PMI_1.2; 1.
Isomerase.
KW isomerase.
SQ SEQUENCE 423 AA; 46601 MW; 90330E53A0F9DBE CRC64;
Query Match 78.9%; Score 45; DB 6; Length 423;
Best Local Similarity 66.7%; Pred. No. 8.3;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ENQDSISSKL 12
Db 77 ENQDSIGSKVD 88
RESULT 6
ID Q7WVJ9 PRELIMINARY; PRT; 74 AA.
AC Q7WVJ9
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL: AF438694; AAP97621.1; -
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8305 MW; CFA43CFDEB8B5B94 CRC64;
Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 ENQDSISSKL 10
Db 3 ENQDSVNSKI 12
RESULT 7
ID Q7WVJ7 PRELIMINARY; PRT; 74 AA.
AC Q7WVJ7
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL: AF438694; AAP97621.1; -
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8305 MW; CFA43CFDEB8B5B94 CRC64;
Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 ENQDSISSKL 10
Db 3 ENQDSVNSKI 12
RESULT 8
ID Q7WV16 PRELIMINARY; PRT; 74 AA.
AC Q7WV16
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL: AF438714; AAP97641.1; -
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8252 MW; C8997CF7E18D139E CRC64;
Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 ENQDSISSKL 10
Db 3 ENQDSVNSKI 12
RESULT 9
ID Q7WVH6 PRELIMINARY; PRT; 74 AA.
AC Q7WVH6
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";

RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438734; AAP97661.1; --
FT NON_TER 1
FT NON_TER 74 74
SQ SEQUENCE 74 AA; 8327 MW; D7B9B56DEB8E7994 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10
DB 3 ENQDSVNSKI 12

RESULT 10

ID Q7WVH4 PRELIMINARY; PRT; 74 AA.
AC Q7WVH4;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the northern Red Sea using phycoerythrin gene sequences."
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438738; AAP97665.1; --
FT NON_TER 1
FT NON_TER 74 74
SQ SEQUENCE 74 AA; 8332 MW; 24A77CFDEB9B02F9 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10
DB 3 ENQDSVNSKI 12

RESULT 11

ID Q7WVH3 PRELIMINARY; PRT; 74 AA.
AC Q7WVH3;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the northern Red Sea using phycoerythrin gene sequences."
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438739; AAP97666.1; --
FT NON_TER 1
FT NON_TER 74 74
SQ SEQUENCE 74 AA; 8362 MW; 27A46DE040E0B2F9 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10
DB 3 ENQDSVNSKI 12

RESULT 12

ID Q7WVG9 PRELIMINARY; PRT; 74 AA.
AC Q7WVG9;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the northern Red Sea using phycoerythrin gene sequences."
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438745; AAP97672.1; --
FT NON_TER 1
FT NON_TER 74 74
SQ SEQUENCE 74 AA; 8294 MW; D7A47CFDEB8E7994 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10
DB 3 ENQDSVNSKI 12

RESULT 13

ID Q7WRK7 PRELIMINARY; PRT; 74 AA.
AC Q7WRK7;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the northern Red Sea using phycoerythrin gene sequences."
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438700; AAP97627.1; --
FT NON_TER 1
FT NON_TER 74 74
SQ SEQUENCE 74 AA; 8308 MW; D7A46E6DEB994994 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred. No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 3 ENQDSVNSKI 12

Search completed: April 19, 2004, 12:00:07
Job time : 4.63158 secs

QY 1 ENQDSISSKL 10
| | | | | :
Db 3 ENQDSVNSKI 12

RESULT 14
Q7WRK6 PRELIMINARY; PRT; 74 AA.
AC Q7WRK6
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus
OX NCEI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438683; AAP97610.1; -.
DR EMBL; AF438722; AAP97649.1; -.
DR EMBL; AF438728; AAP97655.1; -.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8356 MW; D7A4756DEB994994 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred.No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10
| | | | | :
Db 3 ENQDSVNSKI 12

RESULT 15
Q7WRK5 PRELIMINARY; PRT; 74 AA.
AC Q7WRK5
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CpeB (Fragment).
GN CPEB.
OS uncultured Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCEI_TaxID=159733;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22753450; PubMed=12871235;
RA Steglich C., Post A.F., Hess W.R.;
RT "Analysis of natural populations of Prochlorococcus spp. in the
RT northern Red Sea using phycoerythrin gene sequences.";
RL Environ. Microbiol. 5:681-690(2003).
DR EMBL; AF438689; AAP97616.1; -.
DR EMBL; AF438736; AAP97663.1; -.
FT NON_TER 1
FT NON_TER 74
SQ SEQUENCE 74 AA; 8322 MW; D7A47CFDEB994994 CRC64;

Query Match 71.9%; Score 41; DB 2; Length 74;
Best Local Similarity 70.0%; Pred.No. 8;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENQDSISSKL 10

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.07295 Seconds
(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_170_176

Perfect score: 32

Sequence: 1 QAAKAA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCPUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	585	1	US-08-153-799-14
2	32	100.0	585	1	US-08-446-186A-3
3	32	100.0	585	2	US-08-984-176-1
4	32	100.0	585	2	US-08-702-572-2
5	32	100.0	585	3	US-08-769-746-2
6	32	100.0	585	4	US-10-153-064-5
7	32	100.0	609	1	US-08-223-619-3
8	32	100.0	609	1	US-08-433-037-4
9	32	100.0	609	4	US-08-897-956A-2
10	32	100.0	609	4	US-10-153-064-7
11	32	100.0	609	4	US-09-976-594-977
12	32	100.0	609	5	PCT-US95-04075-3
13	32	100.0	610	2	US-08-797-689-2
14	32	100.0	610	4	US-09-984-186-2
15	32	100.0	651	4	US-10-153-064-133
16	32	100.0	652	4	US-10-153-064-96
17	32	100.0	652	4	US-10-153-064-99
18	32	100.0	652	4	US-10-153-064-105
19	32	100.0	652	4	US-10-153-064-132
20	32	100.0	653	4	US-10-153-064-131
21	32	100.0	656	4	US-10-153-064-130
22	32	100.0	660	4	US-10-153-064-90
23	32	100.0	660	4	US-10-153-064-93
24	32	100.0	668	4	US-10-153-064-102
25	32	100.0	676	4	US-10-153-064-95
26	32	100.0	676	4	US-10-153-064-98
27	32	100.0	676	4	US-10-153-064-104

Sequence 127, App
Sequence 129, App
Sequence 125, App
Sequence 123, App
Sequence 92, Appl
Sequence 101, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 3, Appl
Sequence 89, Appl
Sequence 23494, A
Sequence 81, Appl
Sequence 81, Appl
Sequence 81, Appl
Sequence 81, Appl
Sequence 81, Appl

28 32 100.0 676 4 US-10-153-064-127
29 32 100.0 676 4 US-10-153-064-129
30 32 100.0 677 4 US-10-153-064-125
31 32 100.0 680 4 US-10-153-064-123
32 32 100.0 684 4 US-10-153-064-92
33 32 100.0 692 4 US-10-153-064-101
34 32 100.0 783 1 US-08-256-938-2
35 32 100.0 787 1 US-08-256-938-4
36 32 100.0 787 2 US-08-797-689-16
37 32 100.0 787 4 US-08-984-186-16
38 32 100.0 978 4 US-08-897-956A-3
39 32 100.0 1184 4 US-10-153-064-89
40 29 90.6 565 4 US-09-252-991A-23494
41 28 87.5 43 1 US-07-998-003A-81
42 28 87.5 43 1 US-08-453-274B-81
43 28 87.5 43 1 US-08-453-695A-81
44 28 87.5 43 1 US-08-268-161A-81
45 28 87.5 43 2 US-08-453-702A-81

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No 576683
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Scope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 28-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swode, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: Region

LOCATION: 369..419
OTHER INFORMATION: /note= "Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"

FEATURE:
NAME/KEY: Region

LOCATION: 1..585
OTHER INFORMATION: /note= "Amino acid sequence of
OTHER INFORMATION: natural HSA"

US-08-153-799-14

Query Match 100.0%; Score 32; DB 1; Length 585;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7

|||||

Db 170 QAADKAA 176

RESULT 2

US-08-448-196A-3

; Sequence 3, Application US/08448196A

; Patent No. 5780594

; GENERAL INFORMATION:

; APPLICANT: CARTER, DANIEL C.

; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS

; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR

; TITLE OF INVENTION: RELATED PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NASA

; STREET: MARSHALL SPACE FLIGHT CENTER

; CITY: HUNTSVILLE

; STATE: ALABAMA

; COUNTRY: USA

; ZIP: 35812

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/448,196A

; FILING DATE: 23-MAY-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: BROAD JR., ROBERT L.

; REGISTRATION NUMBER: 18,757

; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 205-544-0021

; TELEFAX: 205-544-0258

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 585 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

US-08-448-196A-3

Query Match

Best Local Similarity 100.0%; Score 32; DB 1; Length 585;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7

|||||

Db 170 QAADKAA 176

RESULT 3

US-08-984-176-1

; Sequence 1, Application US/08984176

; Patent No. 5948609

; GENERAL INFORMATION:

; APPLICANT: CARTER, DANIEL C

; APPLICANT: HO, JOSEPH X

; APPLICANT: RUKER, FLORIAN

; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT

; TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

; FILE REFERENCE: 08/984,176

; CURRENT APPLICATION NUMBER: US/08/984,176

; CURRENT FILING DATE: 1997-12-03

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 1

; LENGTH: 585

; TYPE: PRT

; ORGANISM: Homo sapiens

US-08-984-176-1

Query Match

Best Local Similarity 100.0%; Score 32; DB 2; Length 585;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7

|||||

Db 170 QAADKAA 176

RESULT 4

US-08-702-572-2

; Sequence 2, Application US/08702572

; Patent No. 5965386

; GENERAL INFORMATION:

; APPLICANT: Kerry-Williams, Sean M

; APPLICANT: Gilbert, Sarah C

; TITLE OF INVENTION: Yeast Strains and Modified Albumins

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Centeon L.L.C.

; STREET: 1020 First Avenue

; CITY: King of Prussia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19406-1310

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Microsoft Word 6.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/702,572

; FILING DATE: 11-NOV-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO 95/23857

; FILING DATE: 1-MAR-1995

; APPLICATION NUMBER: GB 9404270.2

; FILING DATE: 5-MAR-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Naomi Biswas

; REGISTRATION NUMBER: 38,384

; REFERENCE/DOCKET NUMBER: CE0114 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610/878/4294

; TELEFAX: 610/878/4221

; INFORMATION FOR SEQ ID NO: 2:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2
Query Match 100.0%; Score 32; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2
Query Match 100.0%; Score 32; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5
Query Match 100.0%; Score 32; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 170 QAADKAA 176

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 Beavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3
Query Match 100.0%; Score 32; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 194 QAADKAA 200

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
```


APPLICANT: Barr, Kathryn A.
APPLICANT: Brierley, Russell A.
APPLICANT: Thill, Gregory P.
APPLICANT: Tschopp, Jueki F.
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
TITLE OF INVENTION: PICHIA PASTORIS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/433,037
APPLICATION NUMBER: US/08/433,037
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DiGiglio, Frank S.
REGISTRATION NUMBER: 31,345
REFERENCE/DOCKET NUMBER: 9108Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 100.0%; Score 32; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 194 QAADKAA 200

RESULT 9
US-08-897-956A-2
; Sequence 2, Application US/08897956A
; Patent No. 6423512
; GENERAL INFORMATION:
; APPLICANT: Mary Ellen Digan
; APPLICANT: Philip Lake
; APPLICANT: Hermann Gram
; TITLE OF INVENTION: Fusion polypeptides
; FILE REFERENCE: 600-7244/CPA
; CURRENT APPLICATION NUMBER: US/08/897,956A
; CURRENT FILING DATE: 1997-07-21
; PRIOR APPLICATION NUMBER: 60/022,689
; PRIOR FILING DATE: 1996-07-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 100.0%; Score 32; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 194 QAADKAA 200

RESULT 10
US-10-153-064-7
; Sequence 7, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PFS56
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match 100.0%; Score 32; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 194 QAADKAA 200

RESULT 11
US-09-976-594-977
; Sequence 977, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match 100.0%; Score 32; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 194 QAADKAA 200

RESULT 12
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.

;; APPLICATION NUMBER: PCT/FR93/00085
;; FILING DATE: 28-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith Ph.D., Julie K.
;; REGISTRATION NUMBER: P-38,619
;; REFERENCE/DOCKET NUMBER: ST92006-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (610) 454-3839
;; TELEFAX: (610) 454-3808
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 610 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-797-689-2

Query Match 100.0%; Score 32; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
DB 194 QAADKAA 200
|||||

RESULT 14
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:

;; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
;; TITLE OF INVENTION: Protein
;; NUMBER OF SEQUENCES: 33
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Amgen Center, Patent Operations/RRR
;; STREET: 1840 DeHavilland Drive
;; CITY: Thousand Oaks
;; STATE: California
;; COUNTRY: U.S.
;; ZIP: 91320-1789
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/04075
;; FILING DATE:
;; CLASSIFICATION:
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 609 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
;; PCT-US95-04075-3

Query Match 100.0%; Score 32; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
DB 194 QAADKAA 200
|||||

RESULT 13
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

Query Match 100.0%; Score 32; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
|
|
|
|
|
|
|
Db 194 QAADKAA 200

RESULT 15
US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PFS56
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

Query Match 100.0%; Score 32; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
|
|
|
|
|
|
|
Db 236 QAADKAA 242

Search completed: April 19, 2004, 12:05:19
Job time : 1.07295 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 0.731302 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_247_252

Perfect score: 33

Sequence: 1 HGDLLLE 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	250	2 F83609	Hypothetical prote
2	33	100.0	265	2 I46986	albumin - dog (fra
3	33	100.0	379	2 A57477	potassium channel
4	33	100.0	453	2 A05139	serum albumin - mo
5	33	100.0	600	2 A47391	serum albumin prec
6	33	100.0	605	1 ABPG3	serum albumin prec
7	33	100.0	607	1 ABBS3	serum albumin prec
8	33	100.0	607	1 ABBS3	serum albumin prec
9	33	100.0	607	1 ABBS3	serum albumin prec
10	33	100.0	608	1 ABBS3	serum albumin prec
11	33	100.0	608	2 S57632	serum albumin prec
12	33	100.0	609	1 ABBS3	serum albumin prec
13	33	100.0	609	2 JC5828	albumin - Mongolia
14	33	100.0	714	2 A33595	ATP-dependent DNA
15	31	93.9	132	2 G72774	hypothetical prote
16	31	93.9	159	2 A36600	phosphopantetheine
17	31	93.9	334	2 D37650	pyridoxal phosphat
18	31	93.9	348	2 B75445	conserved hypotet
19	31	93.9	394	2 T32042	hypothetical prote
20	31	93.9	542	2 T08777	probable protein k
21	31	93.9	550	2 H71952	ATP-dependent zinc
22	31	93.9	559	2 A36604	poly(3-hydroxyalka
23	31	93.9	601	2 B81723	conserved hypotet
24	31	93.9	912	1 A53215	protein kinase C (
25	31	93.9	918	1 I48719	protein kinase C (
26	30	90.9	90	2 A05186	hypothetical prote
27	30	90.9	346	2 A95324	hypothetical prote
28	30	90.9	406	2 H84590	hypothetical prote
29	30	90.9	409	2 T25935	hypothetical prote

30 30 90.9 415 2 H84718
31 30 90.9 435 2 F84083
32 30 90.9 495 2 AC3237
33 30 90.9 554 2 G75576
34 30 90.9 560 2 C38604
35 30 90.9 707 2 T02835
36 30 90.9 720 2 T40900
37 30 90.9 816 2 A49151
38 30 90.9 1117 2 AE0075
39 30 90.9 1149 2 T20891
40 30 90.9 1163 2 S07137
41 30 90.9 1271 2 T49009
42 30 90.9 1289 1 RMXRR3
43 30 90.9 1400 1 I38185
44 30 90.9 1513 1 RNR2C2
45 30 90.9 1527 1 RNMZB2

ALIGNMENTS

RESULT 1

F83609
Hypothetical protein PA0278 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83609
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B:
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83609
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-250 <STO>
A:Cross-references: GB:A8004466; GB:A8004091; NID:g9946120; PIDN:AG03667.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0278

Query Match 100.0%; Score 33; DB 2; Length 250;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6

DB 191 HGDLLLE 196

RESULT 2

I46986

albumin - dog (fragment)

C:Species: Canis lupus familiaris (dog)

C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 20-Aug-1999

C:Accession: I46986

R:Spitzauer, S.; Schweiger, C.; Sperr, W.R.; Pandjaitan, B.; Valent, P.; Muhl, S.; Ebner

J. Allergy Clin. Immunol. 93, 614-627, 1994

A:Title: Molecular characterization of dog albumin as a cross-reactive allergen.

A:Reference number: I46986; MUID:94201492; PMID:7512102

A:Accession: I46986

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-265 <SPI>

A:Cross-references: GB:S72946; NID:g633937; PIDN:AAB30434.1; PID:g633938

C:Superfamily: serum albumin; serum albumin repeat homology

F/7-180/Domain: serum albumin repeat homology <SH2>

Query Match 100.0%; Score 33; DB 2; Length 265;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 HGDLLLE 6
DB      57 HGDLLLE 62

RESULT 3
A:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 24-Sep-1999
C:Accession: A57477
R:Watkins, T.; Ishii, T.; Horio, Y.; Morishige, K.I.; Takahashi, N.; Yamada, M.; Yamashita, J. Biol. Chem. 270, 16339-16346, 1995
A:Title: A novel ATP-dependent inward rectifier potassium channel expressed predominantly in the brain
A:Reference number: A57477; MUID:95332346; PMID:7608203
A:Accession: A57477
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-379 <TR>
A:Cross-references: GB:X86818; NID:G939969; PIDN:CAA60501.1; PID:G939970
C:Superfamily: G protein-activated potassium channel protein
C:Keywords: ATP; transmembrane protein

Query Match      100.0%; Score 33; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      93 HGDLLLE 98

RESULT 4
A05139
serum albumin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: A05139; I48638
R:Minighetti, P.P.; Law, S.W.; Dugaiczky, A. Mol. Biol. Evol. 2, 347-358, 1985
A:Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudogenes
A:Reference number: A93055; MUID:88216123; PMID:2452956
A:Accession: A05139
A:Molecule type: mRNA
A:Residues: 1-418 <MIN>
A:Cross-references: GB:M6111; NID:G191764; PIDN:AAA37190.1; PID:G191765
R:Bocaccio, C.; Deschatrette, J.; Meunier-Rotival, M. Gene 88, 181-186, 1990
A:Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the 3' untranslated region of the alpha-fetoprotein gene
A:Accession: I48638
A:Reference number: I48638; MUID:90269606; PMID:1971802
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 379-453 <BOCC>
A:Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F:123-296/Domain: serum albumin repeat homology <SA2>
F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match      100.0%; Score 33; DB 2; Length 453;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      173 HGDLLLE 178

RESULT 5
A47391
serum albumin precursor - rhesus macaque

```

```

C:Species: Macaca mulatta (rhesus macaque)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A47391
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F. Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding
A:Reference number: A47391; MUID:93211971; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-600 <WAT>
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295
A:Experimental source: liver
A:Notes: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)
C:Superfamily: serum albumin; serum albumin repeat homology
F:21-194/Domain: serum albumin repeat homology <SA1>
F:213-386/Domain: serum albumin repeat homology <SA2>
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match      100.0%; Score 33; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      263 HGDLLLE 268

RESULT 6
ABFGS
serum albumin precursor - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S01382; A61006
R:Weinstock, J.; Baldwin, G.S. Nucleic Acids Res. 16, 9045, 1988
A:Title: Nucleotide sequence of porcine liver albumin.
A:Reference number: S01382; MUID:89016582; PMID:3174440
A:Accession: S01382
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-605 <WEI>
A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798
R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M. J. Bone Miner. Res. 4, 235-241, 1989
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral
A:Reference number: A61006; MUID:89269769; PMID:2728927
A:Accession: A61006
A:Molecule type: protein
A:Residues: 23-51, 'X', 53-54, 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XVN', 155 <LIM>
A:Experimental source: dental enamel
A:Note: albumin and other serum proteins are also found in bone
C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
F:17-22/Domain: propeptide #status predicted <PRO>
F:23-605/Product: serum albumin #status predicted <WAT>
F:27-199/Domain: serum albumin repeat homology <SA1>
F:218-391/Domain: serum albumin repeat homology <SA2>
F:410-589/Domain: serum albumin repeat homology <SA3>
F:75-84, 97-113, 112-123, 145-190, 189-198, 221-267, 266-274, 286-300, 299-310, 337-382, 381-390,
F:261/Binding site: bilirubin (Lys) #status predicted

Query Match      100.0%; Score 33; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HGDLLLE 6
DB      268 HGDLLLE 273

```

R;Brown, J.R.
submitted to the Atlas, April 1975
A;Reference number: A94551
A;Accession: A94551
A;Molecule type: protein
A;Residues: 190-195 <BR2>
R;Brown, J.R.
Fed. Proc. 33, 1389, 1974
A;Reference number: A91457
A;Contents: annotation; disulfide bonds
R;Werlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC
A;Reference number: S55232; MUID:95031935; PMID:7945219
A;Accession: S55232
A;Status: preliminary
A;Molecule type: protein
A;Residues: 529-536;569-572 <WER>
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; copper binding; duplication; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-607/Product: serum albumin #status experimental <MPT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4

Query Match 100.0%; Score 33; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
DB 270 HGDLLLE 275

RESULT 8
ABSHS
serum albumin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C;Accession: S06936
R;Brown, W.M.; Driegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A;Reference number: S06936; MUID:90098888; PMID:2602160
A;Accession: S06936
A;Molecule type: mRNA
A;Residues: 1-607 <BRO>
A;Cross-references: EMBL:X17055; NID:g1386; PIDN:CAA34903.1; PID:g1387
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the membra
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-607/Product: serum albumin #status predicted <MAT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F;263/Binding site: bilirubin (lys) #status predicted

Query Match 100.0%; Score 33; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
DB 270 HGDLLLE 275

RESULT 9

ABHOS
serum albumin precursor - horse
C/Species: Equus caballus (domestic horse)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C/Accession: G34053
R/Ho, J.X.; Holowachuk, B.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A/Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A/Reference number: S34053; MUID:93345495; PMID:8344282
A/Accession: G34053
A/Molecule type: mRNA
A/Residues: 1-607 <HOA>
A/Cross-references: GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672
C/Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, steroid hormones (weak bonds with these hormones promote their transfer across the membrane)
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 100.0%; Score 33; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6

DB 270 HGDLLLE 275

RESULT 10

ABRTS
serum albumin precursor - rat
N/Alternate names: preproalbumin
C/Species: Rattus norvegicus (Norway rat)
C/Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 22-Jun-1999
C/Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R/Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A/Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A/Reference number: A93872; MUID:81223722; PMID:7017712
A/Accession: A93872
A/Molecule type: mRNA
A/Residues: 1-608 <SAR>
A/Cross-references: GB:V01222; GB:J00698; NID:G55627; PIDN:CAA24532.1; PID:G55628
R/Stauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A/Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
A/Reference number: A92211; MUID:77249657; PMID:893447
A/Note: Cleavages during protein maturation
A/Accession: A92211
A/Molecule type: protein
A/Residues: 1-38 <STR>
R/Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A/Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A/Reference number: A91946; MUID:78109429; PMID:564345
A/Accession: A91946
A/Molecule type: protein
A/Residues: 25-222 <IS1>
R/Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A/Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino a
A/Reference number: A91940; MUID:76260153; PMID:956149

A/Accession: A91940
A/Molecule type: protein
A/Residues: 223-288;572-608 <IS2>
A/Note: 262-Leu was also found
R/Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A/Title: Copper(II)-binding ability of human alpha-fetoprotein.
A/Reference number: A90758; MUID:79001617; PMID:80265
A/Contents: annotation; copper binding
R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A/Title: Structures of histamine-releasing peptides formed by the action of acid protease
A/Reference number: A45800; MUID:89341406; PMID:2474609
A/Accession: C45800
A/Status: Preliminary
A/Molecule type: protein
A/Residues: 166-173 <CAR>
R/Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A/Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved
A/Reference number: I57621; MUID:87286876; PMID:3475566
A/Accession: I57621
A/Status: Preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-5 <RES>
A/Cross-references: GB:M16825; NID:G202828; PIDN:AAA40712.1; PID:G554412
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status experimental <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-608/Product: serum albumin #status experimental <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status experimental
F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4

Query Match 100.0%; Score 33; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6

DB 271 HGDLLLE 276

RESULT 11

S57632
serum albumin precursor - cat
C/Species: Felis silvestris catus (domestic cat)
C/Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
C/Accession: JC4660; S57632
R/Hilger, C.; Grigoriou, F.; Hentges, F.
Gene 169, 295-296, 1996
A/Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A/Reference number: JC4660; MUID:96194824; PMID:8647469
A/Accession: JC4660
A/Molecule type: mRNA
A/Residues: 1-608 <H12>
A/Cross-references: EMBL:X84842; NID:G886484; PIDN:CAA59279.1; PID:G886485
A/Experimental source: liver
C/Comment: This protein is the major protein component in plasma. It functions as a mult
ein has 35 conserved cysteine residues.
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: liver; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRP>
F:25-608/Product: serum albumin #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match

100.0%; Score 33; DB 2; Length 608;

Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HGDLLE 6
Db 271 HGDLLE 276

RESULT 12
ABHUS
serum albumin precursor [validated] - human
N;Alternate names: preproalbumin
N;Contains: kinetensin
C;Species: Homo sapiens (man)
C;Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000
C;Accession: A93743; I59366; I59286; I59313; G01747; S55314; A91420; S06422; S36
R;Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur
Nucleic Acids Res. 9, 6103-6114, 1981
A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli
A;Reference number: A93743; MUID:82081882; PMID:6171778
A;Accession: A93743
A;Molecule type: mRNA
A;Residues: 1-419; 'K', 421-609 <LAW>
A;Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA23
P;Dugaiczkyk, A.; Law, S.W.; Dennison, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A;Reference number: A93936; MUID:82105994; PMID:6275391
A;Accession: A93936
A;Molecule type: mRNA
A;Residues: 1-120; 'G', 122-609 <DUG>
A;Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590
R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A;Reference number: I39427; MUID:86140099; PMID:2419329
A;Accession: I39427
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-26 <URA>
A;Cross-references: GB:M13075; NID:G178330; PIDN:AAA51688.1; PID:G553173
R;Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.
A;Reference number: I59286; MUID:94181575; PMID:8134387
A;Accession: I59286
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 282-290; 'XSREPLQ' <WAT>
A;Cross-references: GB:S69192; NID:G546032; PIDN:AA30282.1; PID:G546033
A;Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia
R;Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-
A;Reference number: I59313; MUID:94294404; PMID:8022807
A;Accession: I59313
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 589-590; 'ALPRRVKMLIQVLP' <WAD>
A;Cross-references: GB:S70799; NID:G547231; PIDN:AA31177.1; PID:G547232
A;Note: This frame-shift variant is designated albumin Bazzano; four additional variants
R;Menaya, J.; Parrilla, R.; Ayuso, M.S.
Submitted to the EMBL Data Library, March 1995
A;Reference number: G08292
A;Accession: G01747
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-120; 'G', 122-455 <MEN>
A;Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431
R;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
A;Reference number: S55314; MUID:95275251; PMID:7755581

A;Accession: S55314
A;Molecule type: protein
A;Residues: 19-27 <LED>
R;Meloun, B.; Noravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A;Title: Complete amino acid sequence of human serum albumin.
A;Reference number: A91420; MUID:76187907; PMID:1225573
A;Accession: A91420
A;Molecule type: protein
A;Residues: 25-117; 'EQ', 120-154; 'Q', 156-193; 'E', 195-387; 'H', 389-390; 'Y', 392-393; 'A', 395-
R;Roehr, U.; Spittler, G.; Tripiet, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from u
A;Reference number: S06422
A;Note: this paper is in German, with an English abstract
A;Accession: S06422
A;Molecule type: protein
A;Residues: 25-48 <ROB>
R;Pinch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A;Title: Mass spectrometric identification of modifications to human serum albumin treat
A;Reference number: S36882; MUID:93384321; PMID:8373198
A;Accession: S36882
A;Molecule type: protein
A;Residues: 45-67; 141-160; 311-337; 469-490; 570-581 <PIN>
R;Kausler, E.; Spittler, G.
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A;Title: Chem. Hoppe-Seyler 372, 849-855, 1991
A;Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol.
A;Reference number: S17599; MUID:92126241; PMID:1772598
A;Accession: S17599
A;Molecule type: protein
A;Residues: 25-54; 354-357; 431-447 <KAU>
A;Note: 49-Leu was also found
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: A45800
A;Molecule type: protein
A;Residues: 166-173 <CAR>
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; W
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from peptin-tr
A;Reference number: A03239; MUID:86242180; PMID:3087352
A;Accession: A03239
A;Molecule type: protein
A;Residues: 166-173; 'L', <MOG>
R;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A;Title: Mutations in genetic variants of human serum albumin found in Italy.
A;Reference number: A38255; MUID:91062352; PMID:2247440
A;Accession: A38255
A;Molecule type: protein
A;Residues: 76-111 <GAL1>
A;Accession: B38255
A;Molecule type: protein
A;Residues: 82-105; 'K', 107-110 <GAL2>
A;Note: this variant is designated albumin Vibo Valentia
A;Accession: A38255
A;Molecule type: protein
A;Residues: 76-83; 'K', 85-106 <GAL3>
A;Note: this variant is designated albumin Torino
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tanni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin Hei
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: protein
A;Residues: 255-263; 'E', 265-281 <MINI>
A;Note: this variant is designated albumin Herborn
R;Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,
Biochim. Biophys. Acta 1119, 232-238, 1992
A;Title: Two allalbumins with identical electrophoretic mobility are produced by diffe;

A:Reference number: S21078; MUID:92190239; PMID:1347703
A:Accession: S21078
A:Molecule type: protein
A:Residues: 354-356, 'K', 358-378 <MIN2>
A:Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported, R:He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A:Title: Atomic structure and chemistry of human serum albumin.
A:Reference number: A46756; MUID:92334427; PMID:1630489
A:Contents: annotation; X-ray crystallography, 2.8 angstroms
R:Brown, J.R.; Shockley, P.; Behrens, P.O.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,
A:Reference number: A94442
A:Contents: annotation; three-dimensional structure and disulfide bonds
R:Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A:Title: Disulfide bonds in human serum albumin.
A:Reference number: A90930
A:Contents: annotation; disulfide bonds
R:Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A:Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A:Reference number: A90299; MUID:78186630; PMID:656055
A:Contents: annotation; bilirubin-binding site
R:Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20,
A:Title: Serum albumin: conformation and active sites.
A:Reference number: A94408
A:Contents: annotation; binding sites
R:Harper, M.E.; Dugaiczky, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A:Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes
A:Reference number: A90028; MUID:83279982; PMID:6192711
R:Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A:Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.
A:Reference number: A46755; MUID:76257808; PMID:955075
A:Contents: annotation
A:Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid)
R:Bohnev, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 246-248, 1992
A:Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphate
A:Reference number: A56294; MUID:92183881; PMID:1544460
A:Contents: annotation
A:Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in P
atace activity
C:Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
C:Genetics:
A:Gene: GDB:ALB
A:Cross-references: GDB:118990; OMIM:103600
A:Map position: 4q11-4q13
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox
F:19-24/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-609/Product: serum albumin #status experimental <MPT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:166-174/Product: kinetensin #status experimental <KIP>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,148-193,192-201,224-269-277,289-303,302-313,340-385,384-393,4
F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Db 271 HGDLLE 276
RESULT 13
JC5838
albumin - Mongolian jird
C:Species: Meriones unguiculatus (Mongolian jird)
C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 19-May-2000
C:Accession: JC5838
R:Yoshida, K.; Seto-Ohashima, A.; Sinohara, H.
DNA Res. 4, 351-354, 1997
A:Title: Sequencing of cDNA encoding serum albumin and its extrahepatic synthesis in th
A:Reference number: JC5838; MUID:98116663; PMID:9455485
A:Accession: JC5838
A:Molecule type: mRNA
A:Residues: 1-609 <YOS>
A:Cross-references: DDBJ:AB006197; NID:G2317277; PIDN:BAA21765.1; PID:G2317278
A:Experimental source: liver
A:Superfamily: serum albumin; serum albumin repeat homology
F:222-395/Domain: serum albumin repeat homology <SA2>
Query Match 100.0%; Score 33; DB 2; Length 609;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HGDLLE 6
|||||
Db 272 HGDLLE 277
RESULT 14
AE3595
ATP-dependent DNA helicase recg (EC 3.6.1.-) [imported] - Brucella melitensis (strain 1
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AE3595
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Marur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella meliten
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3595
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-714 <KUR>
A:Cross-references: GB:AE080918; PIDN:AAL53928.1; PID:gl7984872; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110586
A:Map position: 11
C:Superfamily: DNA helicase recg
C:Keywords: Hydrolase
Query Match 100.0%; Score 33; DB 2; Length 714;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HGDLLE 6
|||||
Db 666 HGDLLE 671
RESULT 15
G72774
hypothetical protein APE0185 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: G72774
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 93-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A:Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: G72774
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-132 <KAW>
A;Cross-references: DDBJ:AP000058; NID:gs103388; PIDN:BAA79097.1; PID:dl042873; PID:gs103388
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0185
C;Superfamily: Aeropyrum pernix hypothetical protein APE0185

Query Match	93.9%	Score 31;	DB 2;	Length 132;
Best Local Similarity	83.3%	Pred. NO. 18;		
Matches	5;	Conservative	1;	Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLE 6
Db 111 HGDLE 116

Search completed: April 19, 2004, 12:02:27
Job time : 2.7313 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.437673 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_247_252

Perfect score: 33

Sequence: 1 HGDLL6 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	379	1 IRKA HUMAN	P78508 homo sapien
2	33	100.0	379	1 IRKA MOUSE	P4jme3 mus musculus
3	33	100.0	379	1 IRKA RAT	P49655 r atp-sensi
4	33	100.0	600	1 ALBU MACMU	Q28532 macaca mula
5	33	100.0	605	1 ALBU FIG	P08835 sus scrofa
6	33	100.0	607	1 ALBU BOVIN	P02769 bos taurus
7	33	100.0	607	1 ALBU HORSE	P35747 equus caball
8	33	100.0	607	1 ALBU SHEEP	P14639 ovis aries
9	33	100.0	608	1 ALBU CANFA	P49822 canis famil
10	33	100.0	608	1 ALBU FELCA	P49064 felis silve
11	33	100.0	608	1 ALBU MOUSE	P07724 mus musculus
12	33	100.0	608	1 ALBU RABIT	P49065 oryctolagus
13	33	100.0	608	1 ALBU RAT	P02770 rattus norv
14	33	100.0	609	1 ALBU HUMAN	P02768 homo sapien
15	33	100.0	609	1 ALBU MERUN	O35090 meriones un
16	33	100.0	941	1 CDAC HUMAN	Q9un75 homo sapien
17	33	100.0	948	1 CDAC HUMAN	Q9y5h9 homo sapien
18	33	100.0	948	1 CDAA HUMAN	Q9y5i2 homo sapien
19	33	100.0	949	1 CDAB HUMAN	Q9y5i1 homo sapien
20	33	100.0	950	1 CDAB HUMAN	Q9y5h8 homo sapien
21	33	100.0	950	1 CDAD HUMAN	Q9y5i0 homo sapien
22	31	93.9	159	1 COAD PSEAE	Q916d1 pseudomonas
23	31	93.9	280	1 XRC2 HUMAN	Q43543 homo sapien
24	31	93.9	314	1 PRVA CLOTE	Q892r2 clostridium
25	31	93.9	334	1 PDVA BACHD	Q9rc88 bacillus ha
26	31	93.9	393	1 SET8 HUMAN	Q9nqrl homo sapien
27	31	93.9	559	1 PHAA PSEOL	P25494 pseudomonas
28	31	93.9	601	1 Y248 CHLMU	Q9p160 chlamydia m
29	31	93.9	878	1 KPCC HUMAN	Q9b2l6 homo sapien
30	31	93.9	890	1 KPCM HUMAN	Q94806 homo sapien
31	31	93.9	912	1 KPCM HUMAN	Q15139 homo sapien
32	31	93.9	918	1 KPCM MOUSE	Q62101 mus musculus
33	30	90.9	159	1 COAD_PSESM	Q88ah3 pseudomonas

ALIGNMENTS

RESULT 1

ID	IRKA HUMAN	STANDARD;	PRT;	379 AA.
AC	P78508; Q92808;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	ATP-sensitive inward rectifier potassium channel 10 (Potassium channel, inwardly rectifying, subfamily J, member 10) (Inward rectifier K+ channel Kir4.1) (ATP-dependent inwardly rectifying potassium channel Kir4.1).			
DE	rectifier K+ channel Kir4.1.			
DE	potassium channel Kir4.1.			
GN	KCNJ10.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEDLINE=97150765; PubMed=8995301;			
RA	Shuck M.E., Piser T.M., Bock J.H., Slightom J.L., Lee K.S., Bienkowski M.J.;			
RT	"Cloning and characterization of two K+ inward rectifier (Kir) 1.1 potassium channel homologs from human kidney (Kir1.2 and Kir1.3)."; J. Biol. Chem. 272:586-593(1997).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Cerebellum;			
RA	Schoots O., van Tol H.H.M.;			
RL	Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.			
CC	-1- FUNCTION: MAY BE RESPONSIBLE FOR POTASSIUM SUFFERING ACTION OF GLIAL CELLS IN THE BRAIN. INWARD RECTIFIER K(+) CHANNELS ARE CHARACTERIZED BY A GREATER TENDANCY TO ALLOW POTASSIUM TO FLOW INTO THE CELL RATHER THAN OUT OF IT. THEIR VOLTAGE DEPENDANCE IS REGULATED BY THE CONCENTRATION OF EXTRACELLULAR POTASSIUM; AS EXTERNAL K(+) IS RAISED, THE VOLTAGE RANGE OF THE CHANNEL OPENING SHIFTS TO MORE POSITIVE VOLTAGES. THE INWARD RECTIFICATION IS MAINLY DUE TO THE BLOCKAGE OF OUTWARD CURRENT BY INTERNAL MAGNESIUM. CAN BE BLOCKED BY EXTRACELLULAR BA(2+) AND CS(+) (BY SIMILARITY).			
CC	-1- SUBUNIT: SEEMS TO FORM HETERODIMER WITH KIR5.1/KCNJ16 (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- SIMILARITY: Belongs to the inward rectifier-type potassium channel family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U73192; AAC50923.1; -.			
DR	EMBL; U73193; AAC50924.1; -.			

Q88cq7 pseudomonas
P26496 pseudomonas
P12227 plium sativ
P11079 reovirus (t
P56764 arabidopsis
Q8thv5 sinapis alb
Q9mtm3 oenothera h
P38550 nicotiana t
Q04912 homo sapien
P12093 oryza sativ
P16025 zea mays (m
Q88be2 pseudomonas

34 30 90.9 161 1 COAD_PSEPK
35 30 90.9 560 1 PHAC_PSEOL
36 30 90.9 1163 1 RPOD_PEA
37 30 90.9 1289 1 MCE_REOVD
38 30 90.9 1376 1 RPOD_ARATH
39 30 90.9 1384 1 RPOD_SINAL
40 30 90.9 1386 1 RPOD_OENHO
41 30 90.9 1388 1 RPOD_TOBAC
42 30 90.9 1400 1 RON_HUMAN
43 30 90.9 1513 1 RPOD_ORYSA
44 30 90.9 1527 1 RPOD_MAIZE
45 29 87.9 214 1 KGUA_PSESM

DR EMBL; U52155; AAB07046.1; --.
 DR Genew; HGNC:6256; KCNJ10.
 DR MIM; 602208; --.
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:0015272; P: ATP-activated inward rectifier potassium ch. .; TAS.
 DR GO; GO:0006813; P: potassium ion transport; TAS.
 DR InterPro; IPR001838; K-channel_IR.
 DR InterPro; IPR001622; K-channel_pore.
 DR Pfam; PF01007; IRK; 1.
 DR PRINTS; PR01320; KIRCHANNEL.
 DR ProDom; PD001103; K-channel_IR; 2.
 KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
 KW ATP-binding; Potassium transport.
 FT DOMAIN 1 69 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 70 92 M1 (POTENTIAL).
 FT DOMAIN 118 134 H5 (PORE-FORMING) (POTENTIAL).
 FT TRANSMEM 143 167 M2 (POTENTIAL).
 FT DOMAIN 168 379 CYTOPLASMIC (POTENTIAL).
 FT SITE 158 158 ROLE IN THE CONTROL OF POLYAMINE-MEDIATED CHANNEL GATING AND IN THE BLOCKING BY INTRACELLULAR MAGNESIUM (BY SIMILARITY).
 FT NP_BIND 210 217 ATP (POTENTIAL).
 FT CONFLICT 50 50 L -> P (IN REF. 2).
 FT CONFLICT 166 166 L -> Q (IN REF. 2).
 FT CONFLICT 271 271 R -> C (IN REF. 2).
 SQ SEQUENCE 379 AA; 42508 MW; D9DA013FF4003533 CRC64;
 Query Match 100.0%; Score 33; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGDLLLE 6
 DB 93 HGDLLLE 98
 RESULT 2
 IRKA_MOUSE STANDARD; PRT; 379 AA.
 ID IRKA_MOUSE AC Q9JMK3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ATP-sensitive inward rectifier potassium channel 10 (Potassium channel, inwardly rectifying, subfamily J, member 10) (Inward rectifier K+ channel Kir4.1).
 DE rectifier K+ channel Kir4.1).
 GN KCNJ10.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP INANOB A., Takahashi K., Tanemoto M., Fujita A., Kurachi Y.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RL
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1, and ICR; TISSUE=Brain;
 RX MEDLINE=21100968; PubMed=11169792;
 RA Li L., Head V., Timpe L.C.;
 RT "Identification of an inward rectifier potassium channel gene expressed in mouse cortical astrocytes."
 RL Glia 33:57-71 (2001).
 CC -1- FUNCTION: MAY BE RESPONSIBLE FOR POTASSIUM BUFFERING ACTION OF GLIAL CELLS IN THE BRAIN. INWARD RECTIFIER K(+) CHANNELS ARE CHARACTERIZED BY A GREATER TENDENCY TO ALLOW POTASSIUM TO FLOW INTO THE CELL RATHER THAN OUT OF IT. THEIR VOLTAGE DEFENDANCE IS REGULATED BY THE CONCENTRATION OF EXTRACELLULAR POTASSIUM; AS EXTERNAL K(+) IS RAISED, THE VOLTAGE RANGE OF THE CHANNEL OPENING SHIFTS TO MORE POSITIVE VOLTAGES. THE INWARD RECTIFICATION IS MAINLY DUE TO THE BLOCKAGE OF OUTWARD CURRENT BY INTERNAL MAGNESIUM. CAN BE BLOCKED BY EXTRACELLULAR BA(2+) AND CS(+) (BY SIMILARITY).
 CC

CC -1- SURUNIT: SEEMS TO FORM HETERODIMER WITH KIR5.1/KCNJ16 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to the inward rectifier-type potassium channel family.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB039879; BAA92432.1; --.
 DR EMBL; AF322631; AAG42845.1; --.
 DR MGD; MGI:1194504; KCNJ10.
 DR InterPro; IPR001838; K-channel_IR.
 DR InterPro; IPR001622; K-channel_pore.
 DR Pfam; PF01007; IRK; 1.
 DR PRINTS; PR01320; KIRCHANNEL.
 DR ProDom; PD001103; K-channel_IR; 2.
 KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
 KW ATP-binding; Potassium transport.
 FT DOMAIN 1 69 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 70 92 M1 (POTENTIAL).
 FT DOMAIN 118 134 H5 (PORE-FORMING) (POTENTIAL).
 FT TRANSMEM 143 167 M2 (POTENTIAL).
 FT DOMAIN 168 379 CYTOPLASMIC (POTENTIAL).
 FT SITE 158 158 ROLE IN THE CONTROL OF POLYAMINE-MEDIATED CHANNEL GATING AND IN THE BLOCKING BY INTRACELLULAR MAGNESIUM (BY SIMILARITY).
 FT NP_BIND 210 217 ATP (POTENTIAL).
 FT CONFLICT 50 50 L -> P (IN REF. 2).
 FT CONFLICT 166 166 L -> Q (IN REF. 2).
 FT CONFLICT 271 271 R -> C (IN REF. 2).
 SQ SEQUENCE 379 AA; 42432 MW; 7FF08446B7F43453 CRC64;
 Query Match 100.0%; Score 33; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 9.4;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HGDLLLE 6
 DB 93 HGDLLLE 98
 RESULT 3
 IRKA_RAT STANDARD; PRT; 379 AA.
 ID IRKA_RAT AC P49655; Q62790;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ATP-sensitive inward rectifier potassium channel 10 (Potassium channel, inwardly rectifying, subfamily J, member 10) (ATP-sensitive inward rectifier potassium channel KIR4.1) (KIR4.1) (BIR10) (Brain-specific inwardly rectifying K+ channel 1) (BIRK1).
 DE KCNJ10 OR KIR4.2.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=95332346; PubMed=7608203;
 RA Takumi T., Ishii T., Horio Y., Morishige K.-I., Takahashi N., Yamada M., Yamashita T., Kiyama H., Sohmiya K., Nakanishi S., Kurachi Y.;
 RT "A novel ATP-dependent inward rectifier potassium channel expressed predominantly in glial cells."
 RL J. Biol. Chem. 270:16339-16346 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 CC

QY		1	HGDLLLE 6
DB		93	HGDLLE 98
RESULT 4			
ID	ALBU MACMU	STANDARD;	PRT; 600 AA.
AC	Q28522;		
DT	01-NOV-1997	(Rel. 35, Created)	
DD	01-NOV-1997	(Rel. 35, Last sequence update)	
DE	28-FEB-2003	(Rel. 41, last annotation update)	
DE	Serum albumin precursor (Fragment).		
GN	ALB.		
CS	Macaca mulatta (Rhesus macaque).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;		
OC	Cercopithecinae; Macaca.		
OX	NCB1_TaxID=9544;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=93211971; PubMed=8460152;		
RA	Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,		
RA	Dwulet J., Putnam F.W.		
RT	"cDNA and protein sequence of polymorphic macaque albumins that differ		
RT	in bilirubin binding"		
RL	Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).		
CC	-I- FUNCTION: Serum albumin, the main protein of plasma, has a good		
CC	binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,		
CC	hormones, bilirubin and drugs. Its main function is the regulation		
CC	of the colloidal osmotic pressure of blood.		
CC	-I- SUBCELLULAR LOCATION: Secreted.		
CC	-I- TISSUE SPECIFICITY: Plasma.		
CC	-I- SIMILARITY: Belongs to the ALB/AFP/VDB family.		
CC	-I- SIMILARITY: Contains 3 albumin domains.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
DR	EMBL; M90463; AAA36906.1; -		
DR	PIR; A47391; A47391.		
DR	HSP; P02768; IE7B.		
DR	InterPro; IPR00264; Serum albumin.		
DR	Pfam; PF00273; transport_prot; 3.		
DR	PRINTS; PR00802; SERUMALBUMIN.		
DR	ProDom; PD002486; Serum albumin; 1.		
DR	SMART; SM00103; ALBUMIN_3.		
DR	PROSITE; PS00212; ALBUMIN; 3.		
KW	Metal-binding; Lipid-binding; Repeat; Signal; Copper.		
FT	NON TER	1	1
FT	SIGNAL	<1	10
FT	PROPEP	11	16
FT	CHAIN	17	600
FT	DOMAIN	17	197
FT	DOMAIN	204	389
FT	DOMAIN	396	587
FT	METAL	19	19
FT	BINDING	256	256
FT	DISULFID	69	78
FT	DISULFID	91	107
FT	DISULFID	106	117
FT	DISULFID	140	185
FT	DISULFID	194	193
FT	DISULFID	226	262
FT	DISULFID	261	269
FT	DISULFID	281	295
FT	DISULFID	294	305
FT	DISULFID	305	305

FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67980 MW; E45C871A670E740B CRC64;
Query Match 100.0%; Score 33; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HGDLLLE 6
Db 263 HGDLLLE 268
RESULT 5
ALBU_PIG
ID ALBU_PIG STANDARD; PRT; 605 AA.
AC P08835; Q29018; PubMed=3174440;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S.; Weinstock J.;
RT "Nucleotide sequence of porcine liver albumin."
RL Nucleic Acids Res. 16:9045-9045 (1988).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X12422; CAA30970.1; --
CC EMBL; M36787; AAA30988.1; --
CC PIR; S01382; AEPGS.
CC HSP; P02768; IE7H.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; P00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 1
FT PROPEP <1 16 BY SIMILARITY.
FT CHAIN 17 22 BY SIMILARITY.
FT CHAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.

FT METAL 401 592 ALBUMIN 3.
FT DISULFID 31 31 COPPER (BY SIMILARITY).
FT DISULFID 75 84 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1F4FF CRC64;
Query Match 100.0%; Score 33; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HGDLLLE 6
Db 268 HGDLLLE 273
RESULT 6
ALBU_BOVIN
ID ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Holowachuk E.W.; Stoltenberg J.K.; Reed R.G.; Peters T. Jr.;
RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.; AND VARIANT THR-214.
RC TISSUE=Liver;
RA Barry T.; Power S.; Gannon F.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Hilger C.; Grigioni F.; de Beaufort C.; Michel G.; Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.; AND VARIANT THR-214.
RA Wu H.T.; Huang M.C.;
RL "The complete cDNA sequence of bovine serum albumin.";
RN Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-32.
RX MEDLINE=80024278; PubMed=488109;
RA McGilivray R.T.A.; Chung D.W.; Davie E.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-
terminal sequence of prealbumin.";
RL Eur. J. Biochem. 98:477-485 (1979).
RN [6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.

RA Brown J.R.;
RL "Structure of bovine serum albumin.";
RT Fed. Proc. 34:591-591(1975).
RN [7]
RP REVISIONS TO 190-195.
RA Brown J.R.;
RL Submitted (APR-1975) to the PIR data bank.
RN [8]
RP SEQUENCE OF 402-433.
RX MEDLINE=82023364; PubMed=7283978;
RA Reed R.G., Putnam F.W., Peters T. Jr.;
RL "Sequence of residues 400-403 of bovine serum albumin.";
RN Biochem. J. 191:867-868(1980).
RN [9]
RP SEQUENCE OF 19-28.
RX MEDLINE=77134075; PubMed=843354;
RA Patterson J.E., Geller D.M.;
RL "Bovine microsomal albumin: amino terminal sequence of bovine
proalbumin.";
RN Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
RN [10]
RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.
RX MEDLINE=91083649; PubMed=2260975;
RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;
RL "Rapid confirmation and revision of the primary structure of bovine
serum albumin by ESIMS and Frit-FAB LC/MS.";
RN Biochem. Biophys. Res. Commun. 173:639-646(1990).
RN [11]
RP SEQUENCE OF 25-41.
RX MEDLINE=88267456; PubMed=3389500;
RA Hsieh J.C., Lin F.P., Lam M.F.;
RL "Electroblotting onto glass-fiber filter from an analytical
isoelectrofocusing gel: a preparative method for isolating proteins
for N-terminal microsequencing.";
RN Anal. Biochem. 170:1-8(1988).
RN [12]
RP SEQUENCE OF 437-451.
RA Vilbois F.;
RL Submitted (AUG-1998) to Swiss-Prot.
RN [13]
RP DISULFIDE BONDS.
RA Brown J.R.;
RL "Structure of serum albumin: disulfide bridges.";
RN Fed. Proc. 33:1389-1389(1974).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M73993; AAA51411.1; -
DR EMBL; X58989; CAA41735.1; -
DR EMBL; Y17759; CAA76847.1; -
DR EMBL; AF542068; AAN17824.1; -
DR HSSP; P02768; IE7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PRC0802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM01103; ALBUMIN_3.
DR PROSITE; PS00212; ALBUMIN; 3.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
KW Polymorphism. 1 18
FT SIGNAL 19 24
FT PROPEP 25 607
FT CHAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
FT VARIANT 214 214
FT CONFLICT 302 302
FT CONFLICT 304 305
FT CONFLICT 324 324
FT CONFLICT 394 395
FT CONFLICT 437 437
FT CONFLICT 493 494
FT SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;
SQ

Query Match 100.0%; Score 33; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative C; Mismatches 0; Indels 0; Gaps 0;

QY 1 HGDLLLE 6
DB 270 HGDLLLE 275

RESULT 7
ALBU HORSE
ID -ALBU_HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RL "X-ray and primary structure of horse serum albumin (Equus caballus)
at 0.27-nm resolution.";
RN Eur. J. Biochem. 215:205-212(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.

GN ALB.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Beagle; TISSUE=Liver;
 RC Hilger C.; (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RL Submitted (1-24) of dog serum albumin.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=20148667; PubMed=10669848;
 RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H.,
 RA Valenta R., Spitzauer S.;
 RT "Escherichia coli expression and purification of recombinant dog
 albumin, a cross-reactive animal allergen."
 RL J. Allergy Clin. Immunol. 105:279-285(2000).
 RN [3]
 RN SEQUENCE OF 25-48.
 RX MEDLINE=75011422; PubMed=4414013;
 RA Dixon J.W., Sarkar B.;
 RT "Isolation, amino acid sequence and copper(II)-binding properties of
 peptide (1-24) of dog serum albumin."
 RL J. Biol. Chem. 249:5872-5877(1974).
 RN [4]
 RN SEQUENCE OF 25-38.
 RC TISSUE=Heart;
 RX MEDLINE=98163340; PubMed=9504812;
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
 dog heart proteins."
 RL Electrophoresis 18:2795-2802(1997).
 RN [5]
 RN SEQUENCE OF 215-478 FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=94201492; PubMed=7512102;
 RA Spitzauer S., Schweiger C., Speer W.R., Pandjaitan B., Valent P.,
 RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
 RT "Molecular characterization of dog albumin as a cross-reactive
 allergen."
 RL J. Allergy Clin. Immunol. 93:614-627(1994).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- ALLERGEN: Causes an allergic reaction in human.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AJ133489; CAB64867.1; -.
 DR EMBL; Y17737; CAA76841.1; -.
 DR EMBL; S72946; AAB30434.1; -.
 DR HSP; P02768; 1E7B.
 DR HSC-2DPAGE; P49822; DOG.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transprot_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.

FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT DOMAIN 25 205
 FT DOMAIN 212 397
 FT DOMAIN 404 595
 FT METAL 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 148 193
 FT DISULFID 192 201
 FT DISULFID 224 270
 FT DISULFID 269 277
 FT DISULFID 289 303
 FT DISULFID 302 313
 FT DISULFID 340 385
 FT DISULFID 384 393
 FT DISULFID 416 462
 FT DISULFID 461 472
 FT DISULFID 485 501
 FT DISULFID 500 511
 FT DISULFID 538 583
 FT DISULFID 582 591
 FT CONFLICT 1 26
 FT CONFLICT 146 146
 FT CONFLICT 206 206
 FT CONFLICT 349 349
 FT CONFLICT 359 359
 FT CONFLICT 448 448
 FT CONFLICT 474 474
 SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;
 Query Match 100.0%; Score 33; DB 1; Length 608;
 Best Local Similarity 100.0%; Pred. No. 16; Indels 0; Gaps 0;
 Matches 6; Conservative
 QY 1 HGDLLE 6
 DB 271 HGDLLE 276
 RESULT 10
 ID ALBU_FELCA STANDARD; PRT; 608 AA.
 AC P49064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serum albumin precursor (Allergen Fel d 2).
 OS ALB.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=96194824; PubMed=8647469;
 RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
 RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin."
 RL Gene 163:295-296(1996).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- ALLERGEN: Causes an allergic reaction in human.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL; X84842; CAA59279.1; -;
CC PIR; JC4660; S57632.
CC HSSP; P02768; 1E7B.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport Prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SW00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608
FT DOMAIN 25 205
FT DOMAIN 212 397
FT DOMAIN 404 595
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 254 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 593
FT DISULFID 582 591
SQ SEQUENCE 608 AA; 68659 MW; 07B629CAC5F60E5F CRC64;

Query Match 100.0%; Score 33; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 271 HGDLLE 276
1 HGDLLE 6
|||||
271 HGDLLE 276

RESULT 11
ALBU MOUSE STANDARD; PRT; 608 AA.
ID ALBU MOUSE
AC P07724; O61802;
DT 01-APR-1988 (Rel. 07, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB OR ALB1 OR ALB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Tongue;
CC

RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojohori T., Sono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Butt C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatane P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rutterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 99-515 FROM N.A.
RX MEDLINE=89216123; PubMed=2452956;
RA Minghetti P.P., Law S.W., Dugaiczky A.;
RT "The rate of molecular evolution of alpha-fetoprotein approaches that
of pseudogenes.";
RL Mol. Biol. Evol. 2:347-358(1985).
RN [5]
RP SEQUENCE OF 477-551 FROM N.A.
RX STRAIN=BALB/c;
RX MEDLINE=90269606; PubMed=1971802;
RA Boccaccio C., Deschatrette J., Meunier-Rotival M.;
RT "Empty and occupied insertion site of the truncated LINE-1 repeat
located in the mouse serum albumin-encoding gene.";
RL Gene 88:181-186(1990).
RN [6]
RP SEQUENCE OF 25-44.
RC TISSUE=Liver;
RX MEDLINE=93162044; PubMed=1285668;
RA Giometti C.S., Taylor J., Tollaksen S.L.;
RT "Mouse liver protein database: a catalog of proteins detected by two-
dimensional gel electrophoresis.";
RL Electrophoresis 13:970-991(1992).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.

ALBU_RAT
ID ALBU_RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN ALB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=81223722; PubMed=701712;
RX Sargent T.D., Yang M., Bonner J.;
RA "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
[2]
RN SEQUENCE OF 1-38, AND PROCESSING.
RP MEDLINE=77249657; PubMed=893447;
RX Straus A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RA "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
piece. Analysis of the direct translation product of albumin
messenger RNA.";
RL J. Biol. Chem. 252:6846-6855(1977).
[3]
RN SEQUENCE OF 25-222.
RP MEDLINE=78109429; PubMed=564345;
RX Isemura S., Ikenaka T.;
RA "Amino acid sequences of fragments I and II obtained by cyanogen
bromide cleavage of rat serum albumin.";
RL J. Biochem. 83:35-48(1978).
[4]
RN SEQUENCE OF 223-288 AND 572-608.
RP MEDLINE=76260153; PubMed=956149;
RX Isemura S., Ikenaka T.;
RA "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
the amino acid sequences of four fragments.";
RL J. Biochem. 79:1183-1196(1976).
[5]
RN SEQUENCE OF 166-174.
RP TISSUE=Plasma;
RX MEDLINE=87194805; PubMed=2437111;
RA Caraway R.E., Mitra S.P., Cochrane D.E.;
RT "Structure of a biologically active neurotensin-related peptide
obtained from pepsin-treated albumin(s).";
RL J. Biol. Chem. 262:5968-5973(1987).
[6]
RN COPPER-BINDING.
RX MEDLINE=73001517; PubMed=80265;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Copper (II)-binding ability of human alpha-fetoprotein.";
RL Cancer Res. 38:3483-3486(1978).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and
blood flow (potential).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

DR EMBL; V01222; CAA24532.1; --
DR PIR; A93872; ABRTS.
DR HSRF; P02768; 187B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608
FT PEPTIDE 166 174
FT DOMAIN 25 205
FT DOMAIN 212 397
FT DOMAIN 404 595
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
FT VARIANT 262 262
FT CONFLICT 174 174
SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;
Query Match 100.0%; Score 33; DB 1; Length 608;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 HGDLL 6
DB 271 HGDLL 276
RESULT 14
ID ALBU_HUMAN STANDARD; PRT; 609 AA.
AC P02768; O95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=86196112; PubMed=3009475;
RX Minghetti P.P., Ruffer D.E., Kuang W.J., Dennison O.E., Hawkins J.W.;
RA Beattie W.G., Dugaiczyk A.;
RT "Molecular structure of the human albumin gene is revealed by
nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
[2]
RN SEQUENCE FROM N.A., AND VARIANT LYS-420.
RP MEDLINE=82081882; PubMed=6171778;
RX Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,

```
RESULT 8
Q941Y7
ID Q941Y7 PRELIMINARY; PRT; 280 AA.
AC
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE P0431G06.23 protein (B1139B11.1 protein).
GN P0431G06.23 OR B1139B11.1
OS Oryza sativa (rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530, 39947;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T.; Matsumoto T.; Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
clone:P0431G06.23";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T.; Matsumoto T.; Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
clone:B1139B11.1";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AP003683; BAB64711.1; -
DR EMBL; AP004368; BAB90773.1; -
DR Gramene; Q941Y7; -
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR Metal-binding; Zinc; Zinc-finger.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 280 AA; 30370 MW; 097DD3165D61C76 CRC64;

Query Match 90.6%; Score 29; DB 10; Length 280;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 Q94DKAA 7
DB 241 EADKAA 247

RESULT 9
Q8U6W5
ID Q8U6W5 PRELIMINARY; PRT; 311 AA.
AC
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE 2-hydroxyacid dehydrogenase.
GN ATU4691 OR AGR_L_379.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
[1]
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=21608550; PubMed=11743193;
RA Wood D.W.; Setubal J.C.; Kaul R.; Monks D.E.; Kitajima J.P.;
RA Okura V.K.; Zhou Y.; Chen L.; Wood G.E.; Almeida N.F. Jr.; Woo L.;
RA Chen Y.; Paulsen I.T.; Eisen J.A.; Karp P.D.; Bovee D. Sr.;
RA Chapman P.; Clendenning J.; Deatherage G.; Gillet W.; Grant C.;
RA Kutyavina T.; Levy R.; Li M.-J.; McClelland E.; Palmieri A.;
RA Raymond C.; Rouse G.; Saenphimmachak C.; Wu Z.; Romero P.; Gordon D.;
```

```
RA Zhang S.; Yoo H.; Tao Y.; Biddle P.; Jung M.; Krespan W.; Perry M.;
RA Gordon-Kamm B.; Liao L.; Kim S.; Hendrick C.; Zhao Z.-Y.; Dolan M.;
RA Chumley P.; Tingey S.V.; Tomb J.-F.; Gordon M.P.; Olson M.V.;
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58.";
RL Science 294:2317-2323 (2001).
[2]
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=21608551; PubMed=11743194;
RA Goodner B.; Hinkle G.; Gattung S.; Miller N.; Blanchard M.;
RA Quorillo B.; Goldman B.S.; Cao Y.; Askenazi M.; Halling C.; Mullin L.;
RA Houmle K.; Gordon J.; Vaudin M.; Iartchouk O.; Epp A.; Liu F.;
RA Wollam C.; Allinger M.; Doughty D.; Scott C.; Lappas C.; Markelz B.;
RA Flanagan C.; Crowell C.; Gutson J.; Lomo C.; Sear C.; Strub G.;
RA Cielo C.; Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328 (2001).
DR EMBL; AE009397; AAL45485.1; -
DR EMBL; AE008217; AAK88759.1; -
DR PIR; AG3133; AG3133.
DR PIR; E98154; E98154.
DR GO; GO:001616; F:oxidoreductase activity, acting on the CH-O...; IEA.
DR GO; GO:0006564; P:L-serine biosynthesis; IEA.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH_C.
DR Pfam; PF00389; 2-Hacid_DH; 1.
DR Pfam; PF02826; 2-Hacid_DH_C; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 311 AA; 33047 MW; B6C98D520CC962EC CRC64;

Query Match 90.6%; Score 29; DB 16; Length 311;
Best Local Similarity 85.7%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 Q94DKAA 7
DB 30 EADKAA 36

RESULT 10
Q930D0
ID Q930D0 PRELIMINARY; PRT; 315 AA.
AC
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative D-isomer specific 2-hydroxyacid.
GN RA0268 OR SMA0510.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymba (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RC MEDLINE=21396509; PubMed=11481432;
RA Barrett M.J.; Fisher R.F.; Jones T.; Komp C.; Abola A.P.;
RA Barclay-Hubler F.; Bowser L.; Capela D.; Galibert F.; Gouzy J.;
RA Gurjal M.; Hong A.; Huizar L.; Hyman R.W.; Kahn D.; Kahn M.L.;
RA Kallman S.; Keating D.H.; Palm C.; Peck M.C.; Surzycki R.; Wells D.H.;
RA Yeh K.-C.; Davis R.W.; Federspiel N.A.; Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
Sinorhizobium meliloti pSymba megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888 (2001).
DR EMBL; AS007220; AAK64926.1; -
DR PIR; D95295; D95295.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016616; F:oxidoreductase activity, acting on the CH-O...; IEA.
DR GO; GO:0006564; P:L-serine biosynthesis; IEA.
DR InterPro; IPR006139; 2-Hacid_DH.
```

OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=3440;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=2202145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach P.C., Parah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorino C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Canarotte G., Cannavan F., Cardozo J., Chamargo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.H.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Rossi M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A.A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.P., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Seubal J.C., Kitajima J.P.;
 RA "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities";
 RL Nature 417:459-463(2002).
 DR EMBL; AE012229; AA040593.1; --
 DR InterPro; IPR000437; Prok_lipprot_S.
 DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 111 AA; 11336 MW; D2301B0058BC5445 CRC64;

Query Match 90.6%; Score 29; DB 16; Length 111;
 Best Local Similarity 85.7%; Pred. No. 1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 QAADKAA 7
 Db 47 QAADRAA 53
 |||||
 |||||

RESULT 6
 O87592 PRELIMINARY; PRT; 196 AA.
 ID O87592
 AC O87592;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein SC01550.
 GN SC01550 OR SC141.33C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2).
 RA Nagy I., Tamura T., Vanderleyden J., Baumeister W., de Mot R.;
 RT The 20S proteasome of Streptomyces coelicolor.";
 RL J. Bacteriol. 0:0-0(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2).
 RA Seeger K.J., Harris D.;
 RA Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2).
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2).

RX MEDLINE=97000351; PubMed=843436;
 RA Redenbach M., Kieser H.M., Denapait D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper C., Batsman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
 RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AF085832; AAC64274.1; --
 DR EMBL; AL939109; CAB59503.1; --
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 196 AA; 20239 MW; 35E647F0EBE6256C4 CRC64;

Query Match 90.6%; Score 29; DB 16; Length 196;
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 QAADKAA 7
 Db 45 QSAADKAA 51
 |||||
 |||||

RESULT 7
 Q7XPL3 PRELIMINARY; PRT; 274 AA.
 ID Q7XPL3
 AC Q7XPL3;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE OSJNB0085110.13 protein.
 GN OSJNB0085110.13.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 OX NCBI_TaxID=4530;
 [1]
 RP SEQUENCE FROM N.A.
 RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
 RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ping C.W., Sheng H.H.,
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RC EMBL; AL606684; CAE03568.1; --
 SQ SEQUENCE 274 AA; 28866 MW; 499A949503543E5B CRC64;

Query Match 90.6%; Score 29; DB 10; Length 274;
 Best Local Similarity 85.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 QAADKAA 7
 Db 137 EAADKAA 143
 |||||
 |||||

```

Db      450 QAADKAA 456

RESULT 2
Q8YJB3  PRELIMINARY;      PRT;      96 AA.
ID      Q8YJB3;
AC      Q8YJB3;
DT      01-MAR-2002 (TrEMBLrel. 20, Created)
DT      01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Hypothetical cytosolic protein BMEI0173.
GN      BMEI0173.
OS      Brucella melitensis.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Brucellaceae; Brucella.
OX      NCBI_TaxID=29459;
[1]
RN      SEQUENCE FROM N.A.
RC      STRAIN=16M / ATCC 23456 / Biotype 1;
RX      MEDLINE=20020109; PubMed=11756688;
RA      DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA      Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA      Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA      Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA      Haselkorn R., Kyriades N., Overbeek R.;
RA      "The genome sequence of the facultative intracellular pathogen
RT      Brucella melitensis.";
RT      Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR      EMBL; AE009459; ALU51355.1; -;
DR      PIR; AH3273; AH3273.
DR      InterPro; IPR005545; YCII.
DR      Pfam; PF03795; YCII; 1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 96 AA; 10359 MW; 0A876F8F2BCB68F6 CRC64;

Query Match      90.6%; Score 29; DB 16; Length 96;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
Db      56 EAADKAA 62

RESULT 3
Q8FYI4  PRELIMINARY;      PRT;      96 AA.
ID      Q8FYI4;
AC      Q8FYI4;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE      01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE      Conserved hypothetical protein.
GN      BR1890.
OS      Brucella suis.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Brucellaceae; Brucella.
OX      NCBI_TaxID=29461;
[1]
RN      SEQUENCE FROM N.A.
RC      STRAIN=1330 / Bivovar 1;
RX      MEDLINE=22247741; PubMed=12271122;
RA      Paulsen I.T., Seehadi R., Nelson K.E., Eisen J.A., Heidelberg J.P.,
RA      Read T.D., Dodson R.J., Unayam L., Brinkac L.M., Beanan M.J.,
RA      Daugherty S.C., Debey R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA      Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
RA      Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA      Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RA      "The Brucella suis genome reveals fundamental similarities between
RT      animal and plant pathogens and symbionts.";
RT      Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
RL      EMBL; AE014479; AAN30784.1; -;
DR      PIR; AH3273; AH3273.
DR      TIGR; BR1890; -;

Query Match      90.6%; Score 29; DB 16; Length 108;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
Db      45 QAADKAA 51

RESULT 5
Q8PB31  PRELIMINARY;      PRT;      111 AA.
ID      Q8PB31;
AC      Q8PB31;
DT      01-OCT-2002 (TrEMBLrel. 22, Created)
DT      01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein XCC1295.
GN      XCC1295.
OS      Xanthomonas axonopodis (pv. citri).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC      Xanthomonadaceae; Xanthomonas.
OX      NCBI_TaxID=92829;
[1]
RN      SEQUENCE FROM N.A.
RC      STRAIN=306 / ATCC 13902 / XV 101;
RX      MEDLINE=22022145; PubMed=12024217;
RA      da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA      Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA      Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA      Cavarotte G., Cannavan F., Cardoso J., Chamberg J., Ciapina L.P.,
RA      Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA      Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA      Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA      Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA      Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA      Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA      Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA      Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA      Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA      Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA      Setubal J.C., Kitajima J.P.;
RA      "Comparison of the genomes of two Xanthomonas pathogens with differing
RT      host specificities.";
RL      Nature 417:459-463(2002).
DR      EMBL; AE011766; AAM36217.1; -;
DR      InterPro; IPR000437; Prok_lipo_prot_S.
DR      PROSITE; PS00013; PROKAP_LIPOPROTEIN; 1.
KW      Hypothetical protein; Complete proteome.
SQ      SEQUENCE 108 AA; 10946 MW; 9C20BAFDC8574728 CRC64;

Query Match      90.6%; Score 29; DB 16; Length 108;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 QAADKAA 7
Db      45 QAADKAA 51

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 2.70175 Seconds
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_170_176
Perfect score: 32
Sequence: 1 QAADKAA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRMBL_25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacterioph:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	667	Q89BJ7	Q89bj7 bradyrhizob
2	29	90.6	96	Q8FYD8	Q8fyd8 brucella me
3	29	90.6	96	Q8FYD8	Q8fyd8 brucella su
4	29	90.6	108	Q8PMT0	Q8pmt0 xanthomonas
5	29	90.6	111	Q8PBT3	Q8pbt3 xanthomonas
6	29	90.6	196	Q87592	Q87592 streptomyc
7	29	90.6	274	Q7XPL3	Q7xpl3 oryza sativ
8	29	90.6	280	Q941V7	Q941v7 oryza sativ
9	29	90.6	311	Q8U6W5	Q8u6w5 agrobacteri
10	29	90.6	315	Q930D0	Q930d0 rhizobium m
11	29	90.6	387	Q7U0Z1	Q7u0z1 mycobacteri
12	29	90.6	466	Q82PD6	Q82pd6 streptomyc
13	29	90.6	503	Q8NST2	Q8nst2 corynebacte
14	29	90.6	574	Q8XW8	Q8xw8 drosophila
15	29	90.6	584	Q7VSC3	Q7vsc3 felis silve
16	29	90.6	607	Q9W3Z7	Q9w3z7 drosophila

17	29	90.6	906	16	Q8YJG2	Q8yjd2 brucella me
18	29	90.6	906	16	Q8FYD8	Q8fyd8 brucella su
19	29	90.6	1313	16	Q8XRR8	Q8xrr8 raietonia s
20	29	90.6	1713	16	Q82V52	Q82v52 nitrosomona
21	29	90.6	2215	5	Q8IIZ6	Q8iiz6 plasmodium
22	28	87.5	120	16	Q8F6P8	Q8f6p8 leptospira
23	28	87.5	166	2	Q9KJG2	Q9kjg2 porphyromon
24	28	87.5	210	10	Q7XZX2	Q7xxz2 oryza sativ
25	28	87.5	225	17	Q8TUH8	Q8tuh8 methanosarc
26	28	87.5	251	4	Q8N9C7	Q8n9c7 homo sapien
27	28	87.5	260	16	Q911L3	Q911l3 pseudomonas
28	28	87.5	260	16	Q87L38	Q87l38 vibrio para
29	28	87.5	293	16	Q9A2I9	Q9a2i9 caulobacter
30	28	87.5	310	16	Q8A2T3	Q8a2t3 bacteroides
31	28	87.5	320	16	Q9K3G9	Q9k3g9 streptomyc
32	28	87.5	322	6	Q28621	Q28621 oryctolagus
33	28	87.5	325	5	Q26715	Q26715 trypanosoma
34	28	87.5	360	6	Q9K69	Q9k69 oryctolagus
35	28	87.5	371	16	Q7UYV0	Q7uyv0 rhodopirell
36	28	87.5	402	6	Q28622	Q28622 oryctolagus
37	28	87.5	457	16	Q7WQX0	Q7wqx0 bordetella
38	28	87.5	457	16	Q7W1Z2	Q7w1z2 bordetella
39	28	87.5	457	16	Q7VUM3	Q7vum3 bordetella
40	28	87.5	462	16	Q9CHJ4	Q9chj4 lactococcus
41	28	87.5	467	16	Q8EDJ0	Q8edj0 shevanella
42	28	87.5	478	2	Q846S1	Q846s1 myxococcus
43	28	87.5	508	5	Q9N4X2	Q9n4x2 caenorhabdi
44	28	87.5	520	5	Q9BJF4	Q9bjf4 toxoplasma
45	28	87.5	522	5	Q26770	Q26770 trypanosoma

ALIGNMENTS

RESULT 1

Q89BJ7 PRELIMINARY; PRT; 667 AA.
AC Q89BJ7;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE BLR8151 protein.
GN BLR8151.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Ideasa K., Iizuchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005964; BAC53416.1;
DR GO; GO:0003069; F:casease activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001309; ICE_P20.
DR Pfam; PF00656; Peptidase C14; 1.
DR PROSITE; PS0208; CASPASE_P20; 1.
KW Complete proteome.
SQ SEQUENCE 667 AA; 70854 MW; 1394315BD73101C3 CRC64;

Query Match 100.0%; Score 32; DB 16; Length 667;

Best Local Similarity 100.0%; Pred.No. 1.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
|||||

Mon Apr 19 13:27:12 2004

```

RESULT 15
TTF1_CAVPO
ID TTF1_CAVPO STANDARD; PRT; 112 AA.
AC P97273;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thyroid transcription factor 1 (Thyroid nuclear factor 1) (TTF-1)
DE (Homeobox protein Nkx-2.1) (Fragment).
GN TTF1 OR TTF1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunkin-Hartley; TISSUE=Lung;
RA Yuan H.T., Bingle C.D.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Transcription factor that binds and activates the
CC promoter of thyroid specific genes such as thyroglobulin,
CC thyroperoxidase, and thyrotropin receptor. Crucial in the
CC maintenance of the thyroid differentiation phenotype. May play a
CC role in lung development and surfactant homeostasis (By
CC similarity).
CC LOCATION: Nuclear.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the NK-2 homeobox family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U82718; AAB40921.1; -.
DR HSSP; P23441; 1FTT.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Transcription regulation; Activator; Homeobox; DNA-binding;
KW Nuclear protein.
FT NON_TER 1 1
FT DNA_BIND 1 60 HOMEBOX.
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12723 MW; AEEADDF06905F9DB CRC64;

Query Match 84.4%; Score 27; DB 1; Length 112;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 QADKAA 7
Db 59 QAKDKAA 65

```

Search completed: April 19, 2004, 11:52:50
Job time : 1.51062 secs

KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Multifunctional enzyme; FMN; FAD; Flavoprotein; NADP;
 KW Complete proteome. 474
 FT DOMAIN 1 CYTOCHROME P450
 FT DOMAIN 475 NADPH-P-450 REDUCTASE
 FT DOMAIN 493 FLAVOXYDOLIN-LIKE
 FT DOMAIN 666 FAD-BINDING
 FT METAL 403 IRON (HEME AXIAL LIGAND) (BY SIMILARITY)
 SQ SEQUENCE 1061 AA; 119467 MW; 7915DACC20578978 CRC64;

Query Match 87.5%; Score 28; DB 1; Length 1061;
 Best Local Similarity 85.7%; Pred No. 1.7e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 QAAKAA 7
 DB 464 QAAKAA 470

RESULT 13
 DEXT STRDO STANDARD; PRT; 1337 AA.
 AC P39653;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Dextranase precursor (EC 3.2.1.11) (Alpha-1,6-glucan-6-
 DE glucanohydrolase).
 GN DEX.
 OS Streptococcus downei (Streptococcus sobrinus).
 OG Plasmid pYA902.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus
 OX NCBI_TaxID=1317;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-36.
 RC STRAIN=6715 / UAB66;
 RX MEDLINE=94292401; PubMed=8021165;
 RA Wanda S.-Y., Curtiss R. III;
 RT "Purification and characterization of Streptococcus sobrinus
 RT dextranase produced in recombinant Escherichia coli and sequence
 RT analysis of the dextranase gene.";
 RL J. Bacteriol. 176:3839-3850(1994).
 CC -1- FUNCTION: MAY PLAY A ROLE IN SUCROSE-INDEPENDENT ADHERENCE TO THE
 CC PELICLE-COATED TOOTH SURFACE.
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,6-alpha-D-glucosidic
 CC linkages in dextran.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
 CC an amide bond (Potential).
 CC -1- MISCELLANEOUS: The activity of this enzyme is optimal at pH 5.3
 CC and at 39 degrees Celsius.
 CC -1- SIMILARITY: Belongs to family 66 of glycosyl hydrolases.
 CC -1- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M96978; AAA21772.1; -
 CC InterPro; IPR001899; Gram_pos_anchor.
 CC Pfam; PF00746; Gram_pos_anchor; 1.
 CC TIGRfams; TIGR01167; LPXTG_anchor; 1.
 CC PROSITE; PS50847; GRAM_POS_ANCHORING; 1.
 KW Hydrolase; Glycosidase; Cell wall; Peptidoglycan-anchor; Repeat;
 KW Signal; Flasmid.
 FT SIGNAL 1 30
 FT CHAIN 31 1308 DEXTRANASE.
 FT PROPEP 1309 1337 REMOVED BY SORTASE (POTENTIAL).
 FT

FT SITE 1305 1309 LPXTG SORTING SIGNAL (POTENTIAL).
 FT MOD RES 1308 1308 AMIDE-LINKED TO CELL WALL (POTENTIAL).
 SQ SEQUENCE 1337 AA; 143298 MW; B494275A77A2E3D0 CRC64;

Query Match 87.5%; Score 28; DB 1; Length 1337;
 Best Local Similarity 100.0%; Pred No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QAAKAA 6
 DB 968 QAAKAA 973

RESULT 14
 CHIB ARAHY STANDARD; PRT; 46 AA.
 AC Q06013;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Endochitinase IB (EC 3.2.1.14) (CHIT IB) (Fragment).
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionaceae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91094785; PubMed=1980004;
 RA Herget T., Schell J., Schreier P.H.;
 RT "Elicitor-specific induction of one member of the chitinase gene
 RT family in Arachis hypogaea.";
 RL Mol. Gen. Genet. 224:469-476(1990).
 CC -1- FUNCTION: This protein functions as a defense against chitin
 CC containing fungal and bacterial pathogens.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages of N-
 CC acetyl-D-glucosamine polymers of chitin.
 CC -1- INDUCTION: By yeast extract and dilution. Slight induction by
 CC glucan elicitor.
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS IA OR IB (BELONGS TO FAMILY
 CC 19 OF GLYCOSYL HYDROLASES).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X56891; CAA40210.1; -
 CC HSSP; P23951; 2BAA.
 CC InterPro; IPR000726; Glyco_hydro_19.
 CC Pfam; PF00182; Glyco_hydro_19; 1.
 CC PRODOM; PF0354900; Glyco_hydro_19; 1.
 CC PROSITE; PS00773; CHITINASE_19_1; PARTIAL.
 CC PROSITE; PS00774; CHITINASE_19_2; PARTIAL.
 KW Hydrolase; Glycosidase; Chitin degradation; Chitin-binding;
 KW Multigene family.
 FT NON_TER 1 46
 FT NON_TER 46 46
 SQ SEQUENCE 46 AA; 4672 MW; 4F868717729C2B9E CRC64;

Query Match 84.4%; Score 27; DB 1; Length 46;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 AADKAA 7
 DB 22 AADKAA 27

protein synthesis. Protects formylmethionyl-tRNA from spontaneous hydrolysis and promotes its binding to the 30S ribosomal subunits. Also involved in the hydrolysis of GTP during the formation of the 70S ribosomal complex (Cyttoplasmic).

-1- SUBCELLULAR LOCATION: Cytoplasmic.

-1- SIMILARITY: Belongs to the IP-2 family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF017154; AAP96265.1; -
 HAMAP; MF 00100; -; 1.
 InterPro; IPR004161; EFTU D2.
 InterPro; IPR000795; EF_GTPbind.
 InterPro; IPR000178; IP2.
 InterPro; IPR006847; IP2_N_GTP.
 InterPro; IPR005225; Small_GTP.
 Pfam; PF00009; GTP_EFTU; 1.
 Pfam; PF03144; GTP_EFTU D2; 2.
 Pfam; PF04760; IP2_N; 1.
 ProDom; PD186100; IP2; 1.
 TIGRFAMs; TIGR00487; IP-2; 1.
 TIGRFAMs; TIGR00231; small_GTP; 1.
 PROSITE; PS01176; IP2; 1.
 Initiation factor; Protein biosynthesis; GTP-binding;
 Complete proteome.
 DOMAIN 341 490 G-DOMAIN.
 FT NP_BIND 347 354 GTP (BY SIMILARITY).
 FT NP_BIND 394 398 GTP (BY SIMILARITY).
 FT NP_BIND 448 451 GTP (BY SIMILARITY).
 SQ SEQUENCE 839 AA; 91846 MW; BDFP37B8B10461D CRC64;

Query Match 87.5%; Score 28; DB 1; Length 839;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQKAA 7
 |||:|
 65 QAREKAA 71

RESULT 12
 ID CYPD_BACSU STANDARD; PRT; 1061 AA.
 AC O08394;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable bifunctional P-450:NADPH-P450 reductase 1 [Includes:
 DE Cytochrome P450 102 (EC 1.14.14.1); NADPH-cytochrome P450 reductase
 DE (EC 1.6.2.4)]
 GN CYPD OR CYP102A2 OR BSU07250.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97453479; PubMed=9308178;
 RA Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,
 RA Duesterhoeft A., Ehrlich S.D.;
 RT "Sequence of the Bacillus subtilis genome region in the vicinity of
 RT the lev operon reveals two new extracytoplasmic function RNA
 RT polymerase sigma factors SigV and SigZ";
 RL Microbiology 143:2939-2943 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX

 MEDLINE=98044033; PubMed=9384377;
 Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azavedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Enrian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghm S.Y., Glaser P., Golligly E.J., Grandi G.,
 RA Gusepki C., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra F., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viati A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
 RA "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis";
 RL Nature 390:249-256 (1997).
 CC -1- FUNCTION: FUNCTIONS AS A FATTY ACID MONOOXYGENASE. THE REDUCTASE
 CC DOMAIN IS REQUIRED FOR ELECTRON TRANSFER FROM NADP TO CYTOCHROME
 CC P450 (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: NADPH + 2 ferricytochrome = NADP(+) + 2
 CC ferrocycytochrome
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -1- COFACTOR: FAD and FMN (By similarity).
 CC -1- SIMILARITY: In the N-terminal section; belongs to the cytochrome
 CC P450 family.
 CC -1- SIMILARITY: Contains 1 FAD-binding domain.
 CC -1- SIMILARITY: Contains 1 flavodoxin-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 EMBL; D87979; BAA20123.1; -
 DR EMBL; Z99107; CAB12544.1; -
 DR PIR; D69799; D69799.
 DR HSSP; P14779; 1JPZ.
 DR Subtilist; BG12871; cypD.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR003097; FAD_binding.
 DR InterPro; IPR008254; Flav_nitox_synth.
 DR InterPro; IPR001094; Flavodoxin_like.
 DR InterPro; IPR001709; FPN_cyt_reductse.
 DR InterPro; IPR001433; Oxred_FAD/NAD(P).
 DR Pfam; PF00667; FAD_binding_1; 1.
 DR Pfam; PF00358; flavodoxin_1; 1.
 DR Pfam; PF00175; NAD_binding_1; 1.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00369; FLAVODOXIN.
 DR PRINTS; PR00371; FPNCR.
 DR PRINTS; PR00385; P450.
 DR PRINTS; PS00086; CYTOCHROME_P450; 1.
 DR PROSITE; PS00902; FLAVODOXIN_LIKE; 1.
 DR PROSITE; PS00902; FLAVODOXIN_LIKE; 1.
 DR

STCA OR PKSET.
OS *Emicella nidulans* (*Aspergillus nidulans*).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiiales; Trichocomaceae; *Emicella*.
OX NCBI_TaxID=162425;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=FGSC 26;
RC MEDLINE=96202293; PubMed=8643646;
RA Brown D.W., Yu J.-H., Kelkar H.S., Fernandes M., Neebitt T.C.,
RA Keller N.P., Adams T.H., Leonard T.J.;
RT "Twenty-five coregulated transcripts define a sterigmatocystin gene
cluster in *Aspergillus nidulans*."
RT Proc. Natl. Acad. Sci. U.S.A. 93:1418-1422 (1996).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=FGSC 4;
RC MEDLINE=95370159; PubMed=7642507;
RA Yu J.-H., Leonard T.J.;
RT "Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a
novel type I polyketide synthase."
RT J. Bacteriol. 177:4792-4800 (1995).
CC -1- FUNCTION: Involved in the synthesis of the polyketide nucleus of
sterigmatocystin from hexanoyl-CoA and seven malonates.
CC -1- COFACTOR: Contains 2 covalently bound phosphopantetheines
(Potential).
CC -1- PATHWAY: Sterigmatocystin biosynthesis; first step.
CC -1- SIMILARITY: Contains 2 acyl carrier domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; U34740; AAC4191.1; --
DR EMBL; L39121; AAA81586.1; --
DR HSSP; P25715; IMLA.
DR InterPro; IPR008918; 5_3_exo_C.
DR InterPro; IPR001227; AC trans.
DR InterPro; IPR000794; ketoacyl synth.
DR InterPro; IPR006163; Pp bind synth.
DR InterPro; IPR006162; Phosphate S.
DR InterPro; IPR001031; thioesterase.
DR Pfam; PF00698; Acyl_transf; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt_C; 1.
DR Pfam; PF00550; pp-binding; 2.
DR Pfam; PF00975; Thioesterase; 1.
DR PROSITE; PS00075; ACP DOMAIN; 2.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; FALSE NEG.
KW Transferase; Acyltransferase; Phosphopantetheine; Repeat;
KW Multifunctional enzyme.
FT DOMAIN 383 814 BETA-KETOACYL SYNTHASE.
FT DOMAIN 884 1209 ACYL/MALONYL TRANSFERASES.
FT DOMAIN 1706 1777 ACYL CARRIER (ACP) 1.
FT DOMAIN 1830 1901 ACYL CARRIER (ACP) 2.
FT DOMAIN ? 2181 THIOESTERASE.
FT ACT_SITE 552 552 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT_SITE 978 978 ACYL/MALONYL TRANSFERASES (BY
SIMILARITY).
FT BINDING 1738 1738 PHOSPHOPANTHEINE (BY SIMILARITY).
FT BINDING 1862 1862 PHOSPHOPANTHEINE (BY SIMILARITY).
FT ACT_SITE 2028 2028 THIOESTERASE (BY SIMILARITY).
SQ SEQUENCE 2181 AA; 238831 MW; 5A3E5712AA9AD942 CRC64;
Query Match 90.6%; Score 29; DB 1; Length 2181;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 QAAAKAA 7

Db 1637 QASDXAA 1643
|||||
RESULT 10
MOBA THIFE STANDARD; PRT; 409 AA.
AC P22898;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE MOBA protein (Fragment).
GN MOBA.
OS *Thiobacillus ferrooxidans*.
OG Plasmid pTF-FC2.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Acidithiobacillales;
OC Acidithiobacillaceae; Acidithiobacillus.
OX NCBI_TaxID=920;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93015664; PubMed=1400173;
RA Rohrer J., Rawlings D.E.;
RT "Sequence analysis and characterization of the mobilization region of
a broad-host-range plasmid, pTF-FC2, isolated from *Thiobacillus*
ferrooxidans."
RT J. Bacteriol. 174:6230-6237 (1992).
[2]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; M57717; AAA27389.1; --
DR EMBL; A43256; A43256.
DR InterPro; IPR005094; Relaxase.
DR Pfam; PF03432; Relaxase; 1.
KW Plasmid.
FT NON_TER 409 409
SQ SEQUENCE 409 AA; 46836 MW; 27C6593202572FE2 CRC64;
Query Match 87.5%; Score 28; DB 1; Length 409;
Best Local Similarity 85.7%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 QAAAKAA 7
|||||
Db 306 QAAEKAA 312
|||||
RESULT 11
IF2_HAEDU STANDARD; PRT; 839 AA.
AC Q7VLI2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Translation initiation factor IF-2.
GN INF2 OR HD1461
GN *Haemophilus ducreyi*.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; *Haemophilus*.
OX NCBI_TaxID=730;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RT "The complete genome sequence of *Haemophilus ducreyi*."
RT Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: One of the essential components for the initiation of

```

RT complete genome sequence. ";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Usterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains. ";
RL J. Bacteriol. 184:5479-5490 (2002).
CC -!- CATALYTIC ACTIVITY: ATP + succinate + CoA = ADP + succinyl-CoA +
CC phosphate.
CC -!- PATHWAY: Tricarboxylic acid cycle.
CC -!- SUBUNIT: Composed of an alpha chain and a beta chain (By
CC similarity).
CC -!- SIMILARITY: Belongs to the succinate/malate CoA ligase beta
CC subunit family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z79700; CAB01999.1; -.
CC DR EMBL; AE006983; AA045226.1; -.
CC DR PIR; E70716; E70716.
CC DR HSP; P07460; ISCU.
CC DR TIGR; MT0978; -.
CC DR TuberculoList; Rv0951; -.
CC DR HAMAP; MF_00558; -.
CC DR InterPro; IPR003135; ATP-grasp.
CC DR InterPro; IPR005809; CoA_lig_beta.
CC DR InterPro; IPR005811; CoA_ligase.
CC DR Pfam; PF02222; ATP-grasp; 1.
CC DR TIGRFAMs; TIGR01016; succCoAbeta; 1.
CC DR PROSITE; PS01217; SUCCINYL_COA_LIG_3; 1.
CC DR Ligase; tricarboxylic acid cycle; Complete proteome.
CC
CC QUERY MATCH 387 AA; 40925 MW; C3155A7CCAF98FDC CRC64;
CC
CC Query Match 90.6%; Score 29; DB 1; Length 387;
CC Best Local Similarity 85.7%; Pred. No. 40;
CC Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 QAADKAA 7
CC Db 376 EAADKAA 382
CC
CC RESULT 8
CC ALBU_FELCA STANDARD; PRT; 608 AA.
CC ID ALBU_FELCA STANDARD; PRT; 608 AA.
CC AC P49064;
CC DT 01-FEB-1996 (Rel. 33, Created)
CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Serum albumin precursor (Allergen Fel d 2).
CC GN ALB
CC OS Felis silvestris catus (Cat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
CC OX NCBI_TaxID=9685;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=96194824; PubMed=8647469;
CC RA Hilger C., Grigioni F., Kohnen M., Hentges F.;

```

```

RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin. ";
RL Gene 169:295-296 (1996).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X84842; CAAS9279.1; -.
CC DR PIR; JC4660; S57632.
CC DR HSP; P02768; IZ7B.
CC DR InterPro; IPR000264; Serum albumin.
CC DR Pfam; PF00273; transport_prot; 3.
CC DR PRINTS; PR00802; SERUMALBUMIN.
CC DR PRODOM; PD002486; Serum_albumin; 1.
CC DR SMART; SM00103; ALBUMIN; 3.
CC DR PROSITE; PS00213; ALBUMIN; 3.
CC KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
CC
CC METAL-BINDING 18
CC FT SIGNAL 1 18
CC FT PROPEP 19 24
CC FT CHAIN 25 608
CC FT DOMAIN 25 205
CC FT DOMAIN 212 397
CC FT DOMAIN 404 595
CC FT METAL 27 27
CC FT DISULFID 77 86
CC FT DISULFID 99 115
CC FT DISULFID 114 125
CC FT DISULFID 148 193
CC FT DISULFID 192 201
CC FT DISULFID 224 270
CC FT DISULFID 269 277
CC FT DISULFID 289 303
CC FT DISULFID 302 313
CC FT DISULFID 340 385
CC FT DISULFID 384 393
CC FT DISULFID 416 462
CC FT DISULFID 461 472
CC FT DISULFID 485 501
CC FT DISULFID 500 511
CC FT DISULFID 538 583
CC FT DISULFID 582 591
CC FT SEQUENCE 608 AA; 68659 MW; 07E629CAC5F60E5F CRC64;
CC
CC Query Match 90.6%; Score 29; DB 1; Length 608;
CC Best Local Similarity 85.7%; Pred. No. 61;
CC Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 QAADKAA 7
CC Db 194 EAADKAA 200
CC
CC RESULT 9
CC STCA_EMENI STANDARD; PRT; 2181 AA.
CC ID STCA_EMENI STANDARD; PRT; 2181 AA.
CC AC Q12397;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Putative sterigmatocystin biosynthesis polyketide synthase (PKS).

```

```

RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.; with identical electrophoretic mobility are produced
RT "Two albumins with identical electrophoretic mobility are produced

Query Match 100.0%; Score 32; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
DB 194 QAADKAA 200

RESULT 5
RL7_STRVG
ID RL7_STRVG STANDARD; PRT; 126 AA.
AC P4936;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12 (SAL).
GN RPLL.
OS Streptomyces virginiae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1961;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96257210; PubMed=8675024;
RA Katayama M., Sakai Y., Okamoto S., Ihara F., Nihira T., Yamada Y.;
RT "Gene organization in the ada-rplL region of Streptomyces virginiae.";
RL Gene 171:135-136(1996).
CC -1- FUNCTION: Seems to be the binding site for several of the factors
CC involved in protein synthesis and appears to be essential for
CC accurate translation (By similarity).
CC -1- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D50624; BAA09305.1; ALT_INIT.
CC HSP; P02392; ICTF.
CC HAMAP; MF 00368; -; 1.
CC InterPro; IPR008932; Ribos L12/7 olig.
CC InterPro; IPR000206; Ribosomal L12.
CC Pfam; PF00542; Ribosomal L12; 1.
CC ProDom; PD001326; Ribosomal_L12; 1.
CC TIGRFAMs; TIGR00855; L12; 1.
CC INIT MET 0
CC BY SIMILARITY.
CC SEQUENCE 126 AA; 13094 MW; 3A45F0754A65C74B CRC64;

Query Match 90.6%; Score 29; DB 1; Length 126;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
DB 107 EAADKAA 113

RESULT 6
RL7_STRAT
ID RL7_STRAT STANDARD; PRT; 128 AA.
AC P29342;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE 50S ribosomal protein L7/L12.
GN RPLL.
OS Streptomyces antibioticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92380478; PubMed=1511874;
RA Parra F., Blanco G., Alonso J.M., Balbin M., Mendez C., Salas J.A.;
RT "Cloning and sequence of a gene encoding the L7/L12 ribosomal protein
RT equivalent of Streptomyces antibioticus.";
RL Gene 118:127-129(1992).
CC -1- FUNCTION: Seems to be the binding site for several of the factors
CC involved in protein synthesis and appears to be essential for
CC accurate translation (By similarity).
CC -1- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M89911; AAA26811.1; -.
CC HSP; JCI273; JCI273.
CC HSP; P02392; ICTF.
CC HAMAP; MF 00368; -; 1.
CC InterPro; IPR008932; Ribos L12/7 olig.
CC InterPro; IPR000206; Ribosomal L12.
CC Pfam; PF00542; Ribosomal_L12; 1.
CC ProDom; PD001326; Ribosomal_L12; 1.
CC TIGRFAMs; TIGR00855; L12; 1.
CC RIBOSOMAL protein.
CC SEQUENCE 128 AA; 13272 MW; F5C3EE4F45D606E8 CRC64;

Query Match 90.6%; Score 29; DB 1; Length 128;
Best Local Similarity 85.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
DB 109 EAADKAA 115

RESULT 7
SUCC_MYCTU
ID SUCC_MYCTU STANDARD; PRT; 387 AA.
AC P71559;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA synthetase beta chain (EC 6.2.1.5) (SCS-beta).
GN SUCC OR RV0951 OR MT0978 OR MTC10D7.23C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the

```

Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

[7]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Liver, and Skeletal muscle;
 RC MEDLINE=2238257; PubMed=12477532;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toehiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gnaratine P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krywinski M.I., Skalska U., Smalilus D.E.,
 RA Scherich A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SEQUENCE OF 25-609.
 RX MEDLINE=76187907; PubMed=1225573;
 RA Meloun B., Moravsek L., Kostka V.;
 RT "Complete amino acid sequence of human serum albumin.";
 RL FEBS Lett. 58:134-137(1975).
 RN [9]
 RP SEQUENCE OF 25-609.
 RA Brown J.R., Shockley P., Bahrens P.Q.;
 RL (in) Bing D.H. (eds.);
 RL The chemistry and physiology of the human plasma proteins, pp.23-40,
 RL Pergamon Press, New York (1979).
 RN [10]
 RP SEQUENCE OF 1-455 FROM N.A.
 RC TISSUE=Liver;
 RA Menaya J., Parrilla R., Ayuso M.S.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE=86140099; PubMed=2419329;
 RA Urano Y., Watanabe K., Sakai M., Tanaoki T.;
 RT "The human albumin gene. Characterization of the 5' and 3' flanking
 RT regions and the polymorphic gene transcripts.";
 RL J. Biol. Chem. 261:3244-3251(1986).
 RN [12]
 RP SEQUENCE OF 222-229.
 RX MEDLINE=76257808; PubMed=955075;
 RA Walker J.E.;
 RT "Lysine residue 199 of human serum albumin is modified by
 RT acetylsalicylic acid.";
 RL FEBS Lett. 66:173-175(1976).
 RN [13]
 RP SEQUENCE OF 25-44 AND 480-499.
 RC TISSUE=Heart;
 RX MEDLINE=95203287; PubMed=7895732;
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 RT 1994.";
 RL Electrophoresis 15:1459-1465(1994).
 RN [14]
 RP DISULFIDE BONDS.
 RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;
 RT "Disulfide bonds in human serum albumin.";
 RL Collect. Czech. Chem. Commun. 42:564-579(1977).
 RN [15]
 RP BILIRUBIN-BINDING SITE.
 RX MEDLINE=78186830; PubMed=656055;
 RA Jacobsen C.;

"Lysine residue 240 of human serum albumin is involved in high-
 affinity binding of bilirubin";
 Biochem. J. 171:453-459(1978).
 RN [16]
 RP VARIANT CANTERBURY ASN-337.
 RX MEDLINE=87157744; PubMed=3828356;
 RA Brennan S.O., Herbert P.;
 RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
 RT domain of serum albumin.";
 RL Biochim. Biophys. Acta 912:191-197(1987).
 RN [17]
 RP VARIANTS NAG-2 AND NAG-3.
 RX MEDLINE=86085523; PubMed=3479777;
 RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
 RA Satoh C., Neel J.V.;
 RT "Amino acid substitutions in inherited albumin variants from
 RT Amerindian and Japanese populations.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
 RN [18]
 RP VARIANTS NAG-1, HIR-1, HIR-2 AND TOCHIGI.
 RX MEDLINE=89345611; PubMed=2762316;
 RA Arai K., Madison J., Huse K., Ishioka N., Satoh C., Fujita M.,
 RA Neel J.V., Sakurabayashi I., Putnam F.W.;
 RT "Point substitutions in Japanese alloalbumins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
 RN [19]
 RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
 RX MEDLINE=90115905; PubMed=2404284;
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;
 RT "Point substitutions in albumin genetic variants from Asia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
 RN [20]
 RP DESCRIPTION OF VARIANT REDHILL.
 RX MEDLINE=90115852; PubMed=2104980;
 RA Brennan S.O., Wyles T., Peach R.J., Donaldson D., George P.M.;
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
 RT human serum albumin whose precursor has an aberrant signal peptidase
 RT cleavage site.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
 RN [21]
 RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
 RX MEDLINE=91062352; PubMed=2247440;
 RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
 RA Watkins S., Putnam F.W.;
 RT "Mutations in genetic variants of human serum albumin found in
 RT Italy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
 RN [22]
 RP VARIANT VENEZIA.
 RX MEDLINE=91296740; PubMed=2068071;
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
 RA Minchiotti L., Putnam F.W.;
 RT "A donor splice mutation and a single-base deletion produce two
 RT carboxyl-terminal variants of human serum albumin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
 RN [23]
 RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
 RX MEDLINE=92052189; PubMed=1946412;
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
 RA Matsuda Y.-I., Amaki I., Putnam F.W.;
 RT "Genetic variants of serum albumin in Americans and Japanese.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
 RN [24]
 RP VARIANT CASEBROOK ASN-518.
 RX MEDLINE=91316157; PubMed=1859851;
 RA Peach R.J., Brennan S.O.;
 RT "Structural characterization of a glycoprotein variant of human serum
 RT albumin: albumin Casebrook (494 Asp-->Asn).";
 RL Biochim. Biophys. Acta 1097:49-54(1991).
 RN [25]
 RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
 RX MEDLINE=92190239; PubMed=1347703;

"Isolation, amino acid sequence and copper(II)-binding properties of peptide (1-24) of dog serum albumin.",
 J. Biol. Chem. 249:5872-5877(1974).
 [4]
 RN SEQUENCE OF 25-38.
 RC TISSUE=Heart;
 RX MEDLINE=9816340; PubMed=9504812;
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;
 RT "HSC-DPAGE and the two-dimensional gel electrophoresis database of dog heart proteins.",
 RL Electrophoresis 18:2795-2802(1997).
 [5]
 RN SEQUENCE OF 215-478 FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=94201492; PubMed=7512102;
 RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,
 RT "Molecular characterization of dog albumin as a cross-reactive allergen.",
 RL J. Allergy Clin. Immunol. 93:614-627(1994).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- ALLERGEN: Causes an allergic reaction in human.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 DR EMBL; AJ133489; CAB64867.1; .
 DR EMBL; Y17737; CAA76841.1; .
 DR EMBL; S72946; AAB30434.1; .
 DR HSRP; P02768; I57B.
 DR HSC-2DPAGE; P49822; DOG.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transprot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; lipid-binding; Repeat; Signal; Copper; Allergen.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608 SERUM ALBUMIN.
 FT DOMAIN 25 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.
 FT DOMAIN 404 595 ALBUMIN 3.
 FT METAL 27 27 COPPER (BY SIMILARITY).
 FT DISULFID 77 86 BY SIMILARITY.
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 148 193 BY SIMILARITY.
 FT DISULFID 192 201 BY SIMILARITY.
 FT DISULFID 224 270 BY SIMILARITY.
 FT DISULFID 269 277 BY SIMILARITY.
 FT DISULFID 289 303 BY SIMILARITY.
 FT DISULFID 302 313 BY SIMILARITY.
 FT DISULFID 340 385 BY SIMILARITY.
 FT DISULFID 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.
 FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.

FT DISULFID 582 591 BY SIMILARITY.
 FT CONFLICT 1 26 MKWVTFISLFFLFSSAYSGRLVREA -> MDT (IN REF. 2).
 FT CONFLICT 146 146 A -> R (IN REF. 2).
 FT CONFLICT 206 206 I -> T (IN REF. 2).
 FT CONFLICT 349 349 V -> A (IN REF. 2).
 FT CONFLICT 359 359 S -> A (IN REF. 1).
 FT CONFLICT 448 448 V -> VV (IN REF. 5).
 FT CONFLICT 474 474 D -> E (IN REF. 1).
 SQ SEQUENCE 608 AA; 68606 MW; 3CFIC8FF7DD8FC06 CRC64;
 Query Match 100.0%; Score 32; DB 1; Length 608;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 QAAADKAA 7
 Db 194 QAAADKAA 200
 RESULT 4
 ALBU HUMAN STANDARD; PRT: 609 AA.
 ID P02768; Q95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJ20;
 DT 21-JUN-1986 (Rel. 01, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OK NCBI_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86196112; PubMed=3009475;
 RA Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
 RA Beattie W.G., Dugaiczky A.;
 RT "Molecular structure of the human albumin gene is revealed by nucleotide sequence within q11-22 of chromosome 4.",
 RL J. Biol. Chem. 261:6747-6757(1986).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
 RX MEDLINE=82081882; PubMed=6171778;
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
 RA Najarian R.C., Seeburg P.H., Wion K.L.;
 RT "The sequence of human serum albumin cDNA and its expression in E. coli".
 RL Nucleic Acids Res. 9:6103-6114(1981).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
 RX MEDLINE=82105994; PubMed=6275391;
 RA Dugaiczky A., Law S.W., Dennison O.E.;
 RT "Nucleotide sequence and the encoded amino acids of human serum albumin mRNA".
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2615/PRO2675).
 RC TISSUE=Fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
 RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
 RT "Functional prediction of the coding sequences of 121 new genes deduced by analysis of cDNA clones from human fetal liver";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
 RA Huang M.C., Wu H.T.;
 RT "The cDNA sequences of human serum albumin.";


```

FT METAL 19 19 COPPER (BY SIMILARITY)..
FT BINDING 256 256 BILIRUBIN (POTENTIAL)..
FT DISULFID 69 78 BY SIMILARITY..
FT DISULFID 91 107 BY SIMILARITY..
FT DISULFID 106 117 BY SIMILARITY..
FT DISULFID 140 185 BY SIMILARITY..
FT DISULFID 184 193 BY SIMILARITY..
FT DISULFID 216 262 BY SIMILARITY..
FT DISULFID 261 269 BY SIMILARITY..
FT DISULFID 281 295 BY SIMILARITY..
FT DISULFID 294 305 BY SIMILARITY..
FT DISULFID 332 377 BY SIMILARITY..
FT DISULFID 376 385 BY SIMILARITY..
FT DISULFID 408 454 BY SIMILARITY..
FT DISULFID 453 464 BY SIMILARITY..
FT DISULFID 477 493 BY SIMILARITY..
FT DISULFID 492 503 BY SIMILARITY..
FT DISULFID 530 575 BY SIMILARITY..
FT DISULFID 574 583 BY SIMILARITY..
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 100.0%; Score 32; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 186 QAADKAA 192

RESULT 2
ALBU_PIG STANDARD; PRT; 605 AA.
ID ALBU_PIG
AC P08835; Q29018;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016582; PubMed=3174440;
RA Baldwin G.S., Weinstock J.;
RT "Nucleotide sequence of porcine liver albumin.";
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X12422; CAA30970.1; --
CC EMBL; M36787; AAA30988.1; --
CC PIR; S01382; ABPGS.
CC HSSP; P02768; 1E7H.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.

```

```

DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SWART; SMO103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON TER 1 1
FT SIGNAL <1 16 BY SIMILARITY.
FT PROPEP 17 22 BY SIMILARITY.
FT CHAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.
FT DOMAIN 401 592 ALBUMIN 3.
FT METAL 31 31 COPPER (BY SIMILARITY)..
FT DISULFID 75 84 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1F4FF CRC64;

Query Match 100.0%; Score 32; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAADKAA 7
Db 191 QAADKAA 197

RESULT 3
ALBU_CANFA STANDARD; PRT; 608 AA.
ID ALBU_CANFA
AC P49822; O77705; Q9TSZ4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).
GN ALB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RA Hilger C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20148667; PubMed=10669848;
RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H.,
RA Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog
RT albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285(2000).
RN [3]
RP SEQUENCE OF 25-48.
RX MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W., Sarkar B.;

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.510619 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_170_176

Perfect score: 32

Sequence: 1 QAAKAA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	32	100.0	600	1 ALBU_MACMU	Q28522 macaca mula
2	32	100.0	605	1 ALBU_PIG	P08835 sus scrofa
3	32	100.0	608	1 ALBU_CANFA	P49822 canis fami
4	32	100.0	609	1 ALBU_HUMAN	P02768 homo sapien
5	29	90.6	126	1 RL7_STRVG	P48936 streptomyce
6	29	90.6	128	1 RL7_STRAT	P29342 streptomyce
7	29	90.6	387	1 SUCC_MICTO	P71559 mycobacteri
8	29	90.6	608	1 ALBU_FELCA	P49064 felis silve
9	29	90.6	2181	1 STCA_EMENI	Q12397 emerice
10	28	87.5	409	1 MOBA_THIFE	P22898 thiobacilla
11	28	87.5	839	1 IF2_HARDU	Q7V112 haemophilus
12	28	87.5	1061	1 CYPB_BACSU	Q08394 bacillus su
13	28	87.5	1337	1 DEXT_STRDO	P39653 streptococ
14	27	84.4	46	1 CHIB_ARAHY	Q06013 arachis hyp
15	27	84.4	112	1 TTF1_CAVPO	P97273 cavia porce
16	27	84.4	129	1 RS8_ECOLI	P02361 escherichia
17	27	84.4	161	1 HLPB_ECOLI	P11457 escherichia
18	27	84.4	161	1 OMPH_SALTY	P16974 salmonella
19	27	84.4	165	1 OMPH_VERPE	P58607 versinia pe
20	27	84.4	286	1 PTNC_ECOLI	P08187 escherichia
21	27	84.4	266	1 PTRC_KLEFN	P37082 klebsiella
22	27	84.4	238	1 HISI_VIECH	Q9KX44 vibrio chol
23	27	84.4	305	1 ROAO_HUMAN	Q13151 homo sapien
24	27	84.4	353	1 NK24_HUMAN	Q9H224 homo sapien
25	27	84.4	354	1 NK24_MOUSE	Q9eqm3 mus musculu
26	27	84.4	356	1 LC05_HUMAN	Q9ntn7 homo sapien
27	27	84.4	371	1 TTF1_CANFA	P43698 canis fami
28	27	84.4	371	1 TTF1_HUMAN	P43699 homo sapien
29	27	84.4	372	1 TTF1_MOUSE	P50220 mus musculu
30	27	84.4	372	1 TTF1_RAT	P23441 rattus norv
31	27	84.4	378	1 MOBL_THIFE	P20085 thiobacilla
32	27	84.4	384	1 TMPB_TREPH	P29720 treponema p
33	27	84.4	394	1 FLB1_RHIME	Q03842 rhizobium m

ALIGNMENTS

RESULT 1	ID	ALBU_MACMU	STANDARD;	PRT;	600 AA.
AC	Q28522;				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
DE	Serum albumin precursor (fragment).				
GN	ALB				
OS	Macaca mulatta (Rhesus macaque).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;				
OC	Cercopithecoidea; Macaca.				
OX	NCBI_TaxID=9544;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=9321191; PubMed=8460152;				
RA	Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,				
RA	Dwulet J., Putnam F.W.;				
RT	"CDNA and protein sequence of polymorphic macaque albumins that differ				
RT	in bilirubin binding.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).				
CC	-!- FUNCTION: Serum albumin, the main protein of plasma, has a good				
CC	binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,				
CC	hormones, bilirubin and drugs. Its main function is the regulation				
CC	of the colloidal osmotic pressure of blood.				
CC	-!- SUBCELLULAR LOCATION: Secreted.				
CC	-!- TISSUE SPECIFICITY: Plasma.				
CC	-!- SIMILARITY: Belongs to the ALB/AFP/VDB family.				
CC	-!- SIMILARITY: Contains 3 albumin domains.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	EMBL; M90463; AAA36906.1; -				
DR	PIR; A47391; A47391.				
DR	HSSP; P02768; 1E7B.				
DR	InterPro; IPR000264; Serum albumin.				
DR	Pfam; PF00273; transport prot; 3.				
DR	PRINTS; PR00802; SERUMALBUMIN.				
DR	ProDom; PD002486; Serum albumin; 1.				
DR	SMART; SM00103; ALBUMIN; 3.				
DR	PROSITE; PS00212; ALBUMIN; 3.				
KW	Metal-binding; Lipid-binding; Repeat; Signal; Copper.				
FT	NON TER	1			
FT	SIGNAL	<1	10		BY SIMILARITY.
FT	PROPEP	11	16		BY SIMILARITY.
FT	CHAIN	17	600		SERUM ALBUMIN.
FT	DOMAIN	17	197		ALBUMIN 1.
FT	DOMAIN	204	389		ALBUMIN 2.
FT	DOMAIN	396	587		ALBUMIN 3.

Q9ky56 streptomyce
P13119 rhizobium m
P77304 escherichia
P54541 bacillus su
Q8byk5 mus musculu
Q96kr7 homo sapien
P48678 mus musculu
P48679 rattus norv
P52157 streptomyce
Q9kgw3 pseudomonas
P34712 caenorhabdi
Q62504 mus musculu

OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imbraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealsen K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015709; AAN5728.1; -.
DR TIGR; SO2700; -.
DR InterPro; IPR006531; Phage_P2_V.
DR TIGRFAMs; TIGR01644; phage_P2_V; 1.
KW Complete proteome.
SQ SEQUENCE 215 AA; 23197 MW; 8ACFDA22355AD55E CRC64;

Query Match 81.6%; Score 31; DB 16; Length 215;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8
Db |||||
45 ADESADN 51

Search completed: April 19, 2004, 11:59:59
JOB time : 5.08772 secs

Query Match 84.2%; Score 32; DB 5; Length 1128;
Best Local Similarity 75.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 652 VADENAQN 659
|||||

RESULT 12

Q8WQC8 PRELIMINARY; PRT; 1702 AA.
AC Q8WQC8;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T13H2.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RT Science 282:2012-2016(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wu X.;
RT "The sequence of C. elegans cosmid T13H2.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL: U39563; AAL56623.1; -
DR WormPeP: T13H2.4a; C330179.
KW Hypothetical protein.
SQ SEQUENCE 1702 AA; 190617 MW; 1F5CAC942100FEB5 CRC64;

Query Match 84.2%; Score 32; DB 5; Length 1702;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAE 7
Db 129 VADESAE 135
|||||

RESULT 13

Q9F428 PRELIMINARY; PRT; 75 AA.
AC Q9F428;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE TEL1 protein.
GN TEL1.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LO28;
RA Perez-Diaz J.C.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBSJ databases.

DR EMBL: AJ012114; CAC13962.1; -
DR GO: GO:0003677; F:DNA binding; IEA.
DR InterPro: IPR001387; HTH_3.
DR Pfam: PF01381; HTH_3; 1.
DR SMART: SM00530; HTH_XRE; 1.
SQ SEQUENCE 75 AA; 8192 MW; 5F1899B3C945B4BF CRC64;

Query Match 81.6%; Score 31; DB 2; Length 75;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 66 VADDSVEN 73
|||||

RESULT 14

Q8Y784 PRELIMINARY; PRT; 75 AA.
AC Q8Y784;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein Imol411.
GN Imol411.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Suchrieser C., Rueniok C., Amend A.,
Raquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
Charbit A., Chtouani F., Couve E., de Daruvar A., Denoux P.,
Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duesurget O.,
Entian K.-D., Feihl H., Garcia-del Portillo F., Garrido P.,
Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
Rammel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL: AL591979; CAC99489.1; -
DR PIR: AC1251; AC1251.
DR ListList; LMO01411; -
DR GO: GO:0003677; F:DNA binding; IEA.
DR InterPro: IPR001387; HTH_3.
DR SMART: SM00530; HTH_XRE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 75 AA; 8263 MW; 545899B3C549B4E3 CRC64;

Query Match 81.6%; Score 31; DB 16; Length 75;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 66 VADDSVEN 73
|||||

RESULT 15

Q8EDP3 PRELIMINARY; PRT; 215 AA.
AC Q8EDP3;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Prophage MuSo2, baseplate assembly protein V.
GN SO2700.
OS Shewanella oneidensis.

```
RT 60,770 full-length cDNAs.
RL Nature 420:563-573(2002).
DR EMBL; AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Albl.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ NON TER 1
FT SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;

Query Match 84.2%; Score 32; DB 11; Length 576;
Best Local Similarity 87.5%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8
DB 46 VADESAEN 53

RESULT 9
Q8C7H3 PRELIMINARY; PRT; 608 AA.
AC Q8C7H3;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Albumin 1.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI:87991; Albl.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;

Query Match 84.2%; Score 32; DB 11; Length 608;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8
DB 78 VADESAEN 85

RESULT 10
Q8WQC7 PRELIMINARY; PRT; 649 AA.
ID Q8WQC7
AC Q8WQC7;

DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T13H2.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Cottage A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBSJ databases.
DR EMBL; Z81532; CAB04326.3; -.
KW Hypothetical protein.
SQ SEQUENCE 1128 AA; 130372 MW; 6445EE26D7D4775D CRC64;

DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T13H2.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Cottage A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBSJ databases.
DR EMBL; Z81532; CAB04326.3; -.
KW Hypothetical protein.
SQ SEQUENCE 1128 AA; 130372 MW; 6445EE26D7D4775D CRC64;

Query Match 84.2%; Score 32; DB 5; Length 649;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAE 7
DB 129 VADESAE 135

RESULT 11
Q62235 PRELIMINARY; PRT; 1128 AA.
ID Q62235
AC Q62235;
DT 01-AUG-1998 (T-EMBLrel. 07, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein F36F2.3.
GN F36F2.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX Cottage A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBSJ databases.
DR EMBL; Z81532; CAB04326.3; -.
KW Hypothetical protein.
SQ SEQUENCE 1128 AA; 130372 MW; 6445EE26D7D4775D CRC64;
```

DR TIGR; TC0331; --
 DR GO; GO:0003899; F:DNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0016987; F:sigma factor activity; IEA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006352; P:transcription initiation; IEA.
 DR InterPro; IPR009043; RNA_pol_sigma.
 DR InterPro; IPR007627; Sigma70_r2.
 DR InterPro; IPR007624; Sigma70_r3.
 DR InterPro; IPR007630; Sigma70_r4.
 DR InterPro; IPR000943; Sigma70.
 DR Pfam; PF04542; sigma70_r2; 1.
 DR Pfam; PF04539; sigma70_r3; 1.
 DR Pfam; PF04545; sigma70_r4; 1.
 DR PRINTS; PR00046; SIGMA70FC1.
 DR DNA-binding; DNA-directed RNA polymerase; Sigma factor;
 KW Transcription regulation; Complete proteome.
 SQ SEQUENCE 253 AA; 28962 MW; 8365F50D21929538 CRC64;

Query Match 84.2%; Score 32; DB 16; Length 253;
 Best Local Similarity 75.0%; Pred. No. 76;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8
 :|||||
 DB 174 IADERAEN 181

RESULT 6
 Q932D1 PRELIMINARY; PRT; 306 AA.
 AC Q932D1;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE A0501260/PTJ8_240.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
 RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,
 RA Huan V.W., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,
 RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,
 RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis ORF clones."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY057624; AAL15255.1; --
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR002044; CED_4.
 DR Pfam; PF00686; CBM_20; 1.
 DR ProDom; PD001568; CBD_4; 1.
 SQ SEQUENCE 306 AA; 34179 MW; 99E7B50898F407BD CRC64;

Query Match 84.2%; Score 32; DB 10; Length 306;
 Best Local Similarity 75.0%; Pred. No. 93;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
 :|||||
 DB 231 VSDESAQN 238

RESULT 7
 Q9LFB0 PRELIMINARY; PRT; 385 AA.
 AC Q9LFB0;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN F7J8_240.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Zimmermann W., Gruenisen A., Wambutt R., Kalicki J.,
 RA Wohldmann P., Smith A., Bancroft I., Mewes H.W., Rudd S., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL137189; CAB69854.1; --
 DR PIR; T45966; T45966.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR002044; CBD_4.
 DR Pfam; PF00686; CBM_20; 1.
 DR ProDom; PD001568; CBD_4; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 385 AA; 43573 MW; B5BIAC3741836475 CRC64;

Query Match 84.2%; Score 32; DB 10; Length 385;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
 :|||||
 DB 231 VSDESAQN 238

RESULT 8
 Q8C7C7 PRELIMINARY; PRT; 576 AA.
 AC Q8C7C7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Albumin 1 (Fragment).
 GN ALB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6J; TISSUE=Thymus;
 RC MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of

```

ID Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigoidae;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Staderker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RL component in murine infection with Schistosoma mansoni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF418550; AAL08579.1;
DR GO: GO:0005615; C:extracellular space; IEA.
DR GO: GO:0005386; F:carrier activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000264; Serum albumin.
DR Pfam: PF00273; transport; 3.
DR PRINTS: PR002486; Serum albumin; 1.
DR SMART: SM00103; ALBUMIN; 3.
DR PROSITE: PS00212; ALBUMIN; 2.
DR PROSITE: PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; 55EABE28E1C66E54 CRC64;

Query Match 100.0%; Score 38; DB 5; Length 608;
Best Local Similarity 100.0%; Pred. NO. 9.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 78 VADESAEN 85

RESULT 3
ID P94805 PRELIMINARY; PRT; 157 AA.
AC P94805;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Haloflex sp. (strain Aa 2.2).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloferax.
OX NCBI_TaxID=2254;
RN [1]
RP SEQUENCE FROM N.A.
RA Holmes M.L., Dyall-Smith M.L.;
RX MEDLINE=20223653; PubMed=10760168;
RA Holmes M.L., Dyall-Smith M.L.;
RT "Sequence and expression of a halobacterial beta-galactosidase gene.";
RL Mol. Microbiol. 36:114-122 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes M.L., Dyall-Smith M.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: U70664; AAB40124.1; -
DR PIR: T44794; T44794.
DR InterPro: IPR003961; FN III-like.
DR InterPro: IPR008957; FN III-like.
DR Pfam: PF00041; fn3; 1.
DR SMART: SM00060; FN3; 1.
KW Hypothetical protein.
SQ SEQUENCE 157 AA; 17392 MW; 65445110F57407FD CRC64;

Query Match 84.2%; Score 32; DB 1; Length 157;
Best Local Similarity 87.5%; Pred. NO. 45;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 VADESAEN 8
Db 102 VADEIAEN 109

RESULT 4
Q8CG74 PRELIMINARY; PRT; 205 AA.
AC Q8CG74;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Albumin (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=129/SvevTACFB;
RA Van Reeth T., Dreze P.L., Gabant P., Szpirer C., Szpirer J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ277794; CAC81303.1; -
DR GO: GO:0005615; C:extracellular space; IEA.
DR GO: GO:0005386; F:carrier activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000264; Serum albumin.
DR Pfam: PF00273; transport; 1.
DR PRINTS: PR00802; SERUMALBUMIN.
DR SMART: SM00103; ALBUMIN; 1.
DR NON TER 205 205
SQ SEQUENCE 205 AA; 23624 MW; 0A982E85C7474BF9 CRC64;

Query Match 84.2%; Score 32; DB 1; Length 205;
Best Local Similarity 87.5%; Pred. NO. 60;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 78 VADESAEN 85

RESULT 5
Q9PKX9 PRELIMINARY; PRT; 253 AA.
AC Q9PKX9;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE RNA polymerase sigma factor.
GN TC0331.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MoPn/Ni99;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Lither K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406 (2000).
CC -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SIGMA-70 FACTOR FAMILY.
DR EMBL: AE002301; AAF39194.1; -
DR PIR: G81713; G81713.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 3.08772 Seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_54_61

Perfect score: 38

Sequence: 1 VADESAEN 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	396	4 Q81UK7	Q81UK7 homo sapien
2	38	100.0	608	5 Q95VB7	Q95VB7 schistosoma
3	32	84.2	157	1 P94805	P94805 halofexax s
4	32	84.2	205	11 Q8CG74	Q8CG74 mus musculus
5	32	84.2	253	16 Q9PKX9	Q9PKX9 chlamydia m
6	32	84.2	305	10 Q93ZD1	Q93ZD1 arabidopsis
7	32	84.2	385	10 Q9LFB0	Q9LFB0 arabidopsis
8	32	84.2	576	11 Q8C7C7	Q8C7C7 mus musculus
9	32	84.2	608	11 Q8C7H3	Q8C7H3 mus musculus
10	32	84.2	649	5 Q8WQC7	Q8WQC7 caenorhabdi
11	32	84.2	1128	5 Q62235	Q62235 caenorhabdi
12	32	84.2	1702	5 Q8WQC8	Q8WQC8 caenorhabdi
13	31	81.6	75	2 Q9F428	Q9F428 listeria mo
14	31	81.6	75	16 Q8Y784	Q8Y784 listeria mo
15	31	81.6	215	16 Q8EDP3	Q8EDP3 shewanella
16	31	81.6	232	10 Q9MOM3	Q9MOM3 arabidopsis

17	31	81.6	236	16 Q8E684	Q8E684 streptococc
18	31	81.6	236	16 Q8E0L1	Q8E0L1 streptococc
19	31	81.6	274	16 Q7VST1	Q7VST1 bordetella
20	31	81.6	282	16 Q86U2	Q86U2 streptomyce
21	31	81.6	289	2 Q8857	Q8857 streptomyce
22	31	81.6	785	10 Q64756	Q64756 arabidopsis
23	31	81.6	918	10 Q8SA86	Q8SA86 zea mays (m
24	31	81.6	954	10 Q23575	Q23575 arabidopsis
25	31	81.6	1058	10 Q8LP18	Q8LP18 arabidopsis
26	31	81.6	1657	5 Q9XUC3	Q9XUC3 caenorhabdi
27	30	78.9	107	5 Q9VU1	Q9VU1 drosophila
28	30	78.9	109	2 Q04818	Q04818 salmonella
29	30	78.9	189	10 Q93WN1	Q93WN1 arabidopsis
30	30	78.9	219	2 Q9EUB8	Q9EUB8 salmonella
31	30	78.9	221	10 Q82354	Q82354 arabidopsis
32	30	78.9	232	12 Q8QS76	Q8QS76 chimpanzee
33	30	78.9	265	17 Q9HS56	Q9HS56 halobacteri
34	30	78.9	301	2 Q46775	Q46775 escherichia
35	30	78.9	308	2 P71522	P71522 marinococcu
36	30	78.9	309	10 Q9LQY6	Q9LQY6 arabidopsis
37	30	78.9	325	2 Q9S143	Q9S143 escherichia
38	30	78.9	329	16 Q93GR9	Q93GR9 salmonella
39	30	78.9	339	2 P72427	P72427 salmonella
40	30	78.9	345	16 Q7VL69	Q7VL69 haemophilus
41	30	78.9	387	10 Q9SHD8	Q9SHD8 arabidopsis
42	30	78.9	451	5 Q86SK4	Q86SK4 ancylostoma
43	30	78.9	453	5 Q9W140	Q9W140 drosophila
44	30	78.9	466	5 Q93712	Q93712 caenorhabdi
45	30	78.9	475	10 Q9SKT8	Q9SKT8 arabidopsis

ALIGNMENTS

RESULT 1

Q81UK7 PRELIMINARY; PRT; 396 AA.

AC Q81UK7: (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1; -
DR GO; GO:0005635; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; F:transport; IEA.
DR InterPro; IPR00273; transport_prot; 2.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 100.0%; Score 38; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 5.7; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 VADESAEN 8

DB 78 VADESAEN 85

RESULT 2

Q95VB7

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M25787; AAA53283.1; -;
 DR EMBL; M33129; AAA30224.1; -;
 DR EMBL; X52584; CAA36814.1; -;
 DR PIR; A30229; A30229.
 DR InterPro; IPR008882; Trypano PARP.
 DR Pfam; PF05887; Trypano PARP; 1.
 KW Signal; Antigen; Repeat; GPI-anchor; Lipoprotein.
 FT SIGNAL 1 27
 FT CHAIN 28 92 PROCYCLIC FORM SPECIFIC POLYPEPTIDE A-
 FT ALPHA.
 FT PROPEP 93 114
 FT DOMAIN 48 85
 FT LIPID 92 92
 SQ SEQUENCE 114 AA; 11611 MW; FFF2690DAAAE445E CRC64;
 6 X 5 AA TANDEM REPEATS OF G-P-E-T.
 GPI-anchor amidated glycine.

Query Match 76.3%; Score 29; DB 1; Length 114;
 Best Local Similarity 85.7%; Pred. NO. 29;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 ADESAEN 8
 |||||
 Db 28 ADESAEN 34

Search completed: April 19, 2004, 11:52:46
 Job time : 1.58356 secs

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@sb-sib.ch).

CC	EMBL; M23391; AAA28487.1; --.	
CDR	PIR; A30817; A30817.	
CDR	TRANSFAC; T00768; --.	
CDR	FlyBase; FBgnC005642; wdn.	
CDR	InterPro; IPR007087; Znf_C2H2.	
CDR	Pfam; PF00096; zf-C2H2; 7.	
CDR	ProDom; PD000003; Znf_C2H2; 1.	
CDR	SMART; SM00355; Znf_C2H2; 7.	
CDR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 6.	
CDR	PROSITE; PS0157; ZINC_FINGER_C2H2_2; 6.	
KW	Developmental protein; Zinc-finger; Metal-binding; DNA-binding;	
KW	Transcription regulation; Repeat; Nuclear protein.	
FT	DOMAIN 52 85	SER/THR-RICH.
FT	DOMAIN 115 162	SER/THR-RICH.
FT	DOMAIN 141 213	PRO-RICH.
FT	DOMAIN 169 186	GLN-RICH.
FT	ZN_FING 268 292	C2H2-TYPE.
FT	ZN_FING 320 320	C2H2-TYPE.
FT	ZN_FING 330 351	C2H2-TYPE.
FT	ZN_FING 357 379	C2H2-TYPE.
FT	ZN_FING 385 407	C2H2-TYPE.
FT	ZN_FING 413 435	C2H2-TYPE.
FT	ZN_FING 441 463	C2H2-TYPE.
FT	ZN_FING 469 492	C2H2-TYPE.
FT	DOMAIN 521 531	POLY-ALA.
FT	DOMAIN 593 669	PRO-RICH.
FT	DOMAIN 632 647	GLN-RICH.
SO	SEQUENCE 868 AA; 95360 MW; 0234DD17F0EF2BF20 CRC64;	

```

CC -!- SIMILARITY: Contains 1 acyl carrier domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AE014017; AAC027044.1; -.
CC HAMAP, MF_01217; -, 1.
CC InterPro, IPR003231; Acyl carrier.
CC InterPro, IPR006163; Pp_bind.
CC InterPro, IPR006162; Ppantne S.
CC Pfam, PFO0550; pp-binding; 1.
CC TIGRfams, TIGR00517; acyl carrier; 1.
CC PROSITE, PS50075; ACP DOMAIN; 1.
CC PROSITE, PS00012; PHOSPHOPANTETHEINE; 1.
CC Lipid synthesis; Fatty acid biosynthesis; Phosphopantetheine;
CC Complete proteome.
CC FT BINDING 37 37 PHOSPHOPANTETHEINE (BY SIMILARITY).
CC SEQUENCE 78 AA; 9031 MW; CAD574DE83690FD0 CRC64;
CC -----
CC Query Match 76.3%; Score 29; DB 1; Length 78;
CC Best Local Similarity 62.5%; Pred. No. 20;
CC Matches 5; Conservative 2; Indels 0; Gaps 0;

```

```

Query Match          78.9%; Score 30; DB 1; Length 868;
Best Local Similarity 75.0%; Pred.No: 1.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 VADESAEN 8
      |  |  |  |  |
      |  |  |  |  |
DB      90 VLDESAQN 97

RESULT 14
ACBP_BUCBP
ID_ACP_BUCBP STANDARD; PRT; 78 AA.
AC P59449;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE ACyl carrier protein (ACP).
GN ACPP OR BBP322.
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
CC NCBI_taxid=135842;
CX [1]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22426901; PubMed=1252265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
RT 'Reductive genome evolution in Buchnera aphidicola.'
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -I- FUNCTION: Carrier of the growing fatty acid chain in fatty acid
      biosynthesis (By similarity).
CC -I- PATHWAY: De novo fatty acid biosynthesis.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- PTM: 4'-phosphopantetheine is transferred from CoA to a specific
      serine of apo-ACP by acps. This modification is essential for
      activity because fatty acids are bound in thioester linkage to the
      sulphydryl of the prosthetic group (By similarity).

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; AB006197; BAA21765.1; --
CC PIR; JC5838; JC5838.
CC HSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3. Repeat; Signal; Copper.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 609 SERUM ALBUMIN.
FT DOMAIN 25 206 ALBUMIN 1.
FT DOMAIN 213 388 ALBUMIN 2.
FT DOMAIN 405 596 ALBUMIN 3.
FT METAL 28 28 COPPER.
FT METAL 78 87 BY SIMILARITY.
FT DISULFID 100 116 BY SIMILARITY.
FT DISULFID 115 126 BY SIMILARITY.
FT DISULFID 149 194 BY SIMILARITY.
FT DISULFID 193 202 BY SIMILARITY.
FT DISULFID 225 271 BY SIMILARITY.
FT DISULFID 270 278 BY SIMILARITY.
FT DISULFID 290 304 BY SIMILARITY.
FT DISULFID 303 314 BY SIMILARITY.
FT DISULFID 341 386 BY SIMILARITY.
FT DISULFID 385 394 BY SIMILARITY.
FT DISULFID 417 463 BY SIMILARITY.
FT DISULFID 462 473 BY SIMILARITY.
FT DISULFID 486 502 BY SIMILARITY.
FT DISULFID 501 512 BY SIMILARITY.
FT DISULFID 539 584 BY SIMILARITY.
FT DISULFID 583 592 BY SIMILARITY.
SQ SEQUENCE 609 AA; 9CA5F97F67EF1A48 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DESAEN 8
|||
DB 81 DESAEN 86

RESULT 12
SEC9 YEAST
ID SEC9 YEAST STANDARD; PRT; 651 AA.
AC P40357;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein transport protein SEC9.
GN SEC9 OR HSS7 OR YGR009C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95042722; PubMed=7954793;
RA Brennwald P., Kearns B., Champion K., Keraenen S.,
RA Bankaitis V., Novick P.;
RT "Sec9 is a SNAP-25-like component of a yeast SNARE complex that may
be the effector of Sec4 function in exocytosis."

Cell 79:245-258(1994).
[2]
RN SEQUENCE FROM N.A.
RP Hebling U., Hofmann B., Delli H.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF A SNARE COMPLEX THAT MAY BE THE EFFECTOR OF
CC SEC4 FUNCTION IN EXOCYTOSIS.
CC -!- SIMILARITY: Belongs to the SNAP-25 family.
CC -!- SIMILARITY: TO YEAST YMR017W.
CC -!- SIMILARITY: Contains 2 t-SNARE coiled-coil homology domains.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; L34336; AAA35034.1; --
CC EMBL; Z72794; CAA96992.1; --
CC PIR; A55100; A55100.
CC GerOnline; 141321; --
CC SGD; S0003241; SEC9.
CC InterPro; IPR000727; T-SNARE.
CC SMART; SM00397; t-SNARE; 2.
CC PROSITE; PS0192; T-SNARE; 2.
KW Protein transport; Repeat; Coiled coil.
FT DOMAIN 434 496 T-SNARE COILED-COIL HOMOMOLOGY 1.
FT DOMAIN 588 650 T-SNARE COILED-COIL HOMOMOLOGY 2.
SQ SEQUENCE 651 AA; 73623 MW; EA314D73D20A10C7 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 651;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8
|||
DB 335 ADKSAEN 341

RESULT 13
SYCV DROME
ID SYCV DROME STANDARD; PRT; 868 AA.
AC P15619;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serendipity locus protein H-1 (Wings-down protein).
GN SY-C OR WDN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=89039875; PubMed=3141791;
RA Vincent A., Kejzarova-Lepesant J., Segalat L., Yanicostas C.,
RA Lepesant J.-A.;
RT "sry h-1, a new Drosophila melanogaster multifingered protein gene
showing maternal and zygotic expression."
RL Mol. Cell. Biol. 8:4459-4468(1988).
CC -!- FUNCTION: THIS MATERNAL AND ZYGOTIC FINGER PROTEIN MAY BELONG TO
CC A COMPLEX SET OF MULTIFINGERED PROTEINS WHICH PLAY AN IMPORTANT
CC ROLE IN GENE ACTIVATION OR REGULATION AT EARLY EMBRYONIC STAGES
CC THROUGH A MAXIMAL ACCUMULATION OF THEIR TRANSCRIPTS (OR PROTEIN
CC PRODUCT) IN THE MATURE OOCYTE.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: Blastoderm specific.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

RN SEQUENCE FROM N.A.
 RP MEDLINE=97382443; PubMed=9240447;
 RX Zeng Z., Kyaw H., Gakenheimer K.R., Augustus M., Fan P., Zhang X.,
 RA Su K., Carter K.C., Li Y.;
 RT "Cloning, mapping, and tissue distribution of a human homologue of the
 RA mouse jerky gene product.";
 RL Biochem. Biophys. Res. Commun. 236:389-395(1997).
 CC -!- TISSUE SPECIFICITY: Abundantly expressed in the majority of
 CC tissues examined, including brain and skeletal muscle.
 CC -!- SIMILARITY: Contains 1 CENPB domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF004715; AAC65933.1; -;
 DR F1R; JCS594; JCS594.
 DR Genew; HGNC:6200; JRXU.
 DR MTM; 603211; -;
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0007417; P:central nervous system development; TAS.
 DR InterPro; IPR004875; CENP-B.
 DR InterPro; IPR006600; CENPB.
 DR Pfam; PF03184; DDE; 1.
 DR SMART; SMO0674; CENPB; 1.
 FT DOMAIN 104 314 CENP-B.
 FT SEQUENCE 442 AA; 50710 MW; 5EE5430FC54CE488 CRC64;
 SQ
 Query Match 81.6%; Score 31; DB 1; Length 442;
 Best Local Similarity 85.7%; Pred. No. 42;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ADESAEN 8
 DB 375 ADESEN 381
 |||||
 SEQUENCE FROM N.A.
 SYD_XANAC STANDARD; PRT; 478 AA.
 AC Q8PLH2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
 DE (HISRS).
 GN HISS OR XAC1826.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=92829;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vicorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Chapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.M., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,

RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RL host specificities.";
 RL Nature 417:459-463(2002).
 CC -!- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
 CC diphosphate + L-histidyl-tRNA(His).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB011816; AAM36688.1; -;
 DR HAWAP; MF_00127; -; 1.
 DR InterPro; IPR004154; HGTP anticodon.
 DR InterPro; IPR004516; HISS.
 DR InterPro; IPR002314; tRNA-synt 2b.
 DR InterPro; IPR006195; tRNA ligase II.
 DR Pfam; PF03129; HGTP anticodon; 1.
 DR Pfam; PF00587; tRNA-synt 2b; 1.
 DR TIGRFAMs; TIGR00442; HISS; 1.
 DR PROSITE; PS50862; AA tRNA LIGASE II; 1.
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 478 AA; 52630 MW; B2344B1147990E0E CRC64;
 Query Match 78.9%; Score 30; DB 1; Length 478;
 Best Local Similarity 85.7%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 ADESAEN 8
 DB 79 ADEGAEN 85
 |||||
 RESULT 11
 ALBU_MERUN
 ID ALBU_MERUN STANDARD; PRT; 609 AA.
 AC O35050;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Meriones.
 OX NCBI_TaxID=10047;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGS IDR; TISSUE=Liver;
 RX MEDLINE=98115663; PubMed=9455485;
 RA Yoshida K., Seto-Oshihama A., Sinohara H.;
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
 RL synthesis in the Mongolian gerbil, Meriones unguiculatus.";
 CC DNA Res. 4:351-354(1997).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.

DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transpore prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 DR KW Metal-binding; ALBUMIN; 3.
 DR SIGNAL 1 18
 DR FT PROPEP 19 24
 DR FT CHAIN 25 608
 DR FT DOMAIN 25 205
 DR FT DOMAIN 212 397
 DR FT DOMAIN 404 595
 DR FT METAL 27 27
 DR FT DISULFID 77 86
 DR FT DISULFID 99 115
 DR FT DISULFID 114 125
 DR FT DISULFID 148 193
 DR FT DISULFID 192 201
 DR FT DISULFID 224 270
 DR FT DISULFID 269 277
 DR FT DISULFID 289 303
 DR FT DISULFID 302 313
 DR FT DISULFID 340 385
 DR FT DISULFID 384 393
 DR FT DISULFID 416 462
 DR FT DISULFID 461 472
 DR FT DISULFID 485 501
 DR FT DISULFID 500 511
 DR FT DISULFID 538 583
 DR FT DISULFID 582 591
 DR SEQUENCE 608 AA; CF5E92647AAFE9A2 CRC64;
 Query Match 84.2%; Score 32; DB 1; Length 608;
 Best Local Similarity 87.5%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VADESSEN 8
 DB 78 VADESSEN 85
 RESULT 8
 LAG1_YEAST STANDARD; PRT; 411 AA.
 AC P38703; 1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Longevity-assurance protein 1 (Longevity assurance factor 1).
 GN LAG1 OR YHL003C.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RX MEDLINE=94253121; PubMed=8195187;
 RA D'Mello N.P., Childress A.M., Franklin D.S., Kale S.P.,
 RA Pinnwasdi C., Jazwinska S.M.;
 RT "Cloning and characterization of LAG1, a longevity-assurance gene in
 yeast.";
 RL J. Biol. Chem. 269:15451-15459(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L.W., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaubin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 VIII";
 RL Science 265:2077-2082(1994).
 CC -!- FUNCTION: Involved in the aging process. Deletion of LAG1 results
 CC in a pronounced increase (approximately 50%) in mean and in
 CC maximum life span.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- SIMILARITY: Contains 1 TLC (TRAM/LAG1/CLN8) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U08133; AAA21579.1; -;
 DR EMBL; U10555; AAB84429.1; -;
 DR PIR; S46800; S46800.
 DR Germonline; 139270; -;
 DR SGD; S0000995; LAG1.
 DR GO; GO:0005783; C:Endoplasmic reticulum; IDA.
 DR GO; GO:0007574; P:cell aging (sensu Saccharomycetes); IMP.
 DR GO; GO:0046513; P:ceramide biosynthesis; IMP.
 DR InterPro; IPR005547; LAG1.
 DR InterPro; IPR006634; TLC.
 DR Pfam; PF03798; LAG1; 1.
 DR SMART; SM00724; TLC; 1.
 DR PROSITE; PS00922; TLC; 1.
 KW Transmembrane.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 135 155 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 212 232 POTENTIAL.
 FT TRANSMEM 252 272 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 356 376 POTENTIAL.
 FT DOMAIN 168 384 TLC.
 FT CONFLICT 173 174 ML -> IV (IN REF. 2).
 FT CONFLICT 220 220 F -> C (IN REF. 2).
 FT CONFLICT 301 411 VTFPFVGLVFFVFIYLRHVVVNIRILSVLTERHGNVYL
 NPATQQYKWSIPIVFVLIQAQLVNLVFLILLYRL
 IWQIQKDSRSDSDSEAESEKCE -> TEISGIWE
 KOIDSNDNPTRELSNETSKQVPLVLPNTENRAL
 LEAIKSRVPTAIIDTDEPSLVTPPGNDLSLVNPLL
 GVLRAGQSGQLQNLARNNEK (IN REF. 1).
 SQ SEQUENCE 411 AA; 48454 MW; 91676D58AC053F3C CRC64;
 Query Match 81.6%; Score 31; DB 1; Length 411;
 Best Local Similarity 85.7%; Pred. No. 39;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ADESAEN 8
 DB 397 SDESAEN 403
 RESULT 9
 JRKL_HUMAN
 ID JRKL_HUMAN STANDARD; PRT; 442 AA.
 AC Q9Y4A0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Jerky protein homolog like (HHMJG).
 GN JRKL
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Locuallano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton B., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grumwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 99-516 FROM N.A.
RX MEDLINE=88216123; PubMed=2452956;
RA Minchetti P.P., Law S.W., Dugaiczky A.,
RA "The rate of molecular evolution of alpha-fetoprotein approaches that
RT of pseudogenes.";
RL Mol. Biol. Evol. 2:347-358 (1985).
RN [5]
RP SEQUENCE OF 477-551 FROM N.A.
RX STRAIN=BALB/C;
RA MEDLINE=90269606; PubMed=1971802;
RA Boccaccio C., Deschattre J., Neunier-Rotival M.,
RA "Empty and occupied insertion site of the truncated LINE-1 repeat
RT located in the mouse serum albumin-encoding gene.";
RL Gene 88:181-186 (1990).
RN [6]
RP SEQUENCE OF 25-44.
RX TISSUE=Liver;
RA MEDLINE=93162044; PubMed=1286668;
RA Glometti C.S., Taylor J., Tollaksen S.L.,
RA "Mouse liver protein database: a catalog of proteins detected by two-
RT dimensional gel electrophoresis.";
RL Electrophoresis 13:970-991 (1992).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ011413; CAA09617.1; -;
CC EMBL; AK010025; BAB26650.1; -;
CC EMBL; BC049971; AAB49971.1; -;
CC EMBL; M16111; AAA37190.1; -;
CC EMBL; X13060; CAA31458.1; -;
CC PIR; A05139; A05139.
CC HSP; P02768; 187B.
CC SWISS-2DFAGE; P07724; MOUSE.
CC MGD; MGI:87991; Albl.
CC InterPro; IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SMO0103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Metal-binding; lipid-binding; Repeat; Signal; Copper.
CC SIGNAL 1 18 BY SIMILARITY.
CC PROPEP 19 24

FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT CONFLICT 27 27 H -> D (IN REF. 6).
FT CONFLICT 33 33 H -> D (IN REF. 6).
FT CONFLICT 41 41 Q -> I (IN REF. 6).
SQ SEQUENCE 608 AA; 292F7C7BED3A61B4 CRC64;
Query Match 84.2%; Score 32; DB 1; Length 608;
Best Local Similarity 87.5%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VADESAEN B
Db 78 VADESAEN 85
RESULT 7
ALBU RABIT
ID ALBU RABIT STANDARD; PRT; 608 AA.
AC P49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U18344; AAB58347.1; -;
CC HSP; P02768; 187B.

```

SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;
Query Match 92.1%; Score 35; DB 1; Length 608;
Best Local Similarity 87.5%; Pred. No. 8;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
DB 78 VADENAEN 85

RESULT 5
ALBU HORSE
ID ALBU HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)
RT at 0.27-nm resolution.";
RL Bur. J. Biochem. 215:205-212(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds Ige.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X74045; CAA52194.1; -
CC F1R; S34053; ABHOS.
CC DR HSP; P02768; 1E7B.
CC DR InterPro; IPR000264; Serum albumin.
CC DR Pfam; PF00273; transport prot; 3.
CC DR PRINTS; PR00802; SERUMALBUMIN.
CC DR ProDom; PD002486; Serum albumin; 1.
CC DR SMART; SM00103; ALBUMIN; 3.
CC DR PROSITE; PS00312; ALBUMIN; 3.
CC DR Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
CC SIGNAL 1 18
CC FT PROPEP 19 24
CC FT CHAIN 25 607
CC FT DOMAIN 25 204
CC FT DOMAIN 211 396
CC FT DOMAIN 403 594
CC FT METAL 27 27
CC FT METAL 77 86
CC FT DISULFID 99 115
CC FT DISULFID 114 125
CC FT DISULFID 147 192
CC FT DISULFID 191 200
CC FT DISULFID 191 200
CC FT DISULFID 223 269

```

```

SQ SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;
Query Match 89.5%; Score 34; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8
DB 79 ADESAEN 85

RESULT 6
ALBU MOUSE
ID ALBU MOUSE STANDARD; PRT; 608 AA.
AC P07724; O61802;
DT 01-APR-1988 (Rel. 07, Created)
DT 15-JUN-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB OR ALB1 OR ALB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayaishizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=23388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Suetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```

RA Winchiotti L., Putnam F.W.;
 RT "A donor splice mutation and a single-base deletion produce two
 RL carboxyl-terminal variants of human serum albumin.";
 RN Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
 [23]
 RN VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
 RP KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
 RX MEDLINE=92052189; PubMed=1946412;
 RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
 RT Matsuda Y.-I., Amaki I., Putnam F.W.;
 RA "Genetic variants of serum albumin in Americans and Japanese.";
 RN Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
 [24]
 RN VARIANT CASEBROOK ASN-518.
 RP MEDLINE=91316157; PubMed=1859851;
 RA Peach R.J., Brennan S.O.;
 RT "Structural characterization of a glycoprotein variant of human serum
 RL albumin: albumin Casebrook (494 Asp->Asn).";
 RN Biochim. Biophys. Acta 1097:49-54(1991).
 [25]
 RN VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
 RP MEDLINE=92190239; PubMed=1347703;
 RA Winchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
 RT Rochu D., Forta F.;
 RA "Two allalbumins with identical electrophoretic mobility are produced
 RT "Two allalbumins with identical electrophoretic mobility are produced
 Query Match 100.0%; Score 38; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 1.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VADESAEN 8
 DB 78 VADESAEN 85

RESULT 4
 ALBU_RAT
 ID ALBU_RAT STANDARD; PRT; 608 AA.
 AC P02770; P11382;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
 GN ALB
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=8123722; PubMed=7017712;
 RA Sargent T.D., Yang M., Bonner J.;
 RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
 RN Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
 [2]
 RN SEQUENCE OF 1-38, AND PROCESSING.
 RP MEDLINE=7724967; PubMed=893447;
 RA Straus A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
 RT "rat liver pre-proalbumin: complete amino acid sequence of the pre-
 piece. Analysis of the direct translation product of albumin
 messenger RNA.";
 RN J. Biol. Chem. 252:6846-6855(1977).
 [3]
 RN SEQUENCE OF 25-222.
 RP MEDLINE=78109429; PubMed=564345;
 RA Isemura S., Ikenaka T.;
 RT "Amino acid sequences of fragments I and II obtained by cyanogen
 bromide cleavage of rat serum albumin.";
 RN J. Biochem. 83:35-48(1978).
 [4]
 RN SEQUENCE OF 223-288 AND 572-608.
 RP MEDLINE=76260153; PubMed=956149;
 RA Isemura S., Ikenaka T.;

"Fragmentation of rat serum albumin by cyanogen bromide cleavage and
 the amino acid sequences of four fragments.";
 J. Biochem. 79:1183-1196(1976).
 [5]
 RN SEQUENCE OF 166-174.
 RP TISSUE=Plasma;
 RX MEDLINE=87194805; PubMed=2437111;
 RA Carraway R.E., Mitra S.P., Cochran D.E.;
 RT "Structure of a biologically active neurotensin-related peptide
 obtained from pepain-treated albumin(s).";
 RN J. Biol. Chem. 262:5968-5973(1987).
 [6]
 RP COPPER-BINDING.
 RX MEDLINE=79001617; PubMed=80265;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";
 RN Cancer Res. 38:3483-3486(1978).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -1- FUNCTION: NRP regulates fat digestion, lipid absorption, and
 blood flow (potential).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.
 CC -1- SIMILARITY: Contains 3 albumin domains.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

 CC EMBL; V01222; CAA24532.1; -;
 DR PIR; A93872; ABRTS.
 DR HSP; P02768; 1E7B.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR PRODOM; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT PEPTIDE 166 174
 FT DOMAIN 25 205
 FT DOMAIN 212 397
 FT DOMAIN 404 595
 FT METAL 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 148 193
 FT DISULFID 192 201
 FT DISULFID 224 270
 FT DISULFID 269 277
 FT DISULFID 289 303
 FT DISULFID 302 313
 FT DISULFID 340 385
 FT DISULFID 384 393
 FT DISULFID 416 462
 FT DISULFID 461 472
 FT DISULFID 485 501
 FT DISULFID 500 511
 FT DISULFID 538 583
 FT DISULFID 582 591
 FT VARIANT 262 262
 FT CONFLICT 174 174
 Y -> L (IN REF. 5).

RX MEDLINE=82105994; PubMed=6275391;
 RA Dugaiczkyk A., Law S.W., Dennison O.E.;
 RT "Nucleotide sequence and the encoded amino acids of human serum
 albumin mRNA";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).
 RC TISSUE=Fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
 XU W., Gao P., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
 RT "Functional prediction of the coding sequences of 121 new genes
 deduced by analysis of cDNA clones from human fetal liver";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
 RA Huang M.C., Wu H.T.;
 RT "The cDNA sequences of human serum albumin";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Liver, and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 Rosa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
 Basak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Vallalón D.K., Muray D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettelman M., Madan A., Rodríguez S., Sanchez A.,
 Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodríguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RN SEQUENCE OF 25-609.
 RX MEDLINE=76187907; PubMed=1225573;
 RA Meloun B., Moravsek L., Kostka V.;
 RT "Complete amino acid sequence of human serum albumin";
 RL FEBS Lett. 58:134-137(1975).
 RN [9]
 RN SEQUENCE OF 25-609.
 RA Brown J.R., Shockley P., Behrens P.Q.;
 RL (in) Bing D.H. (eds.);
 RL The chemistry and physiology of the human plasma proteins, pp.23-40,
 Pergamon Press, New York (1979).
 RN [10]
 RN SEQUENCE OF 1-455 FROM N.A.
 RC TISSUE=Liver;
 RA Menaya J., Farrilla R., Ayuso M.S.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RN SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE=86140099; PubMed=2419329;
 RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
 RT "The human albumin gene. Characterization of the 5' and 3' flanking
 regions and the polymorphic gene transcripts";
 RL J. Biol. Chem. 261:3244-3251(1986).
 RN [12]
 RN SEQUENCE OF 222-229.
 RX MEDLINE=76257808; PubMed=955075;
 RA Walker J.E.;
 RT "Lysine residue 199 of human serum albumin is modified by
 acetylserine sulfonamide";
 RL FEBS Lett. 66:173-175(1976).
 RN [13]
 RN SEQUENCE OF 25-44 AND 480-499.
 RC TISSUE=Heart;
 RX MEDLINE=95203287; PubMed=7895732;
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 1994";
 RL Electrophoresis 15:1459-1465(1994).
 RN [14]
 RN DISULFIDE BONDS.
 RA Saber M.A., Stockbauer P., Moravsek L., Meloun B.;
 RT "Disulfide bonds in human serum albumin";
 RL Collect. Czech. Chem. Commun. 42:564-579(1977).
 RN [15]
 RN BILIRUBIN-BINDING SITE.
 RX MEDLINE=78186630; PubMed=656055;
 RA Jacobsen C.;
 RT "Lysine residue 240 of human serum albumin is involved in high-
 affinity binding of bilirubin";
 RL Biochem. J. 171:453-459(1978).
 RN [16]
 RN VARIANT CANTERBURY ASN-337.
 RX MEDLINE=87157744; PubMed=3828358;
 RA Brennan S.O., Herbert P.;
 RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
 domain of serum albumin";
 RL Biochim. Biophys. Acta 912:191-197(1987).
 RN [17]
 RN VARIANTS NAG-2 AND NAG-3.
 RX MEDLINE=88069523; PubMed=3479777;
 RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
 Satoh C., Neel J.V.;
 RT "Amino acid substitutions in inherited albumin variants from
 Amerindian and Japanese populations";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
 RN [18]
 RN VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
 RX MEDLINE=89345611; PubMed=2762316;
 RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,
 Neel J.V., Sakurabayashi I., Putnam F.W.;
 RT "Point substitutions in Japanese allolalbumins";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
 RN [19]
 RN VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
 RX MEDLINE=90115905; PubMed=2404284;
 RA Arai K., Madison J., Shimizu A., Putnam F.W.;
 RT "Point substitutions in albumin genetic variants from Asia";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
 RN [20]
 RN DESCRIPTION OF VARIANT REDHILL.
 RX MEDLINE=90115952; PubMed=2104980;
 RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
 RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
 human serum albumin whose precursor has an aberrant signal peptidase
 cleavage site";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
 RN [21]
 RN VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
 RX MEDLINE=91062352; PubMed=2247440;
 RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
 Watkins S., Putnam F.W.;
 RT "Mutations in genetic variants of human serum albumin found in
 Italy";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
 RN [22]
 RN VARIANT VENEZIA.
 RX MEDLINE=91296740; PubMed=2068071;
 RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,

```

FT METAL 19 19 COPPER (BY SIMILARITY).
FT BINDING 256 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 375 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; E45C871A670E740B CRC64;

Query Match 100.0%; Score 38; DB 1; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
DB 70 VADESAEN 77

RESULT 2
ALBU_PIG STANDARD; PRT; 605 AA.
AC P08835; Q29018;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Baldwin G.S., Weinstein J.;
RT "Nucleotide sequence of porcine liver albumin.";
RL Nucleic Acids Res. 16:9045-9045(1988).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X12422; CAA30970.1; -
CC EMBL; M36787; AAA30988.1; -
CC PIR; S01382; ABPGS.
CC HSP; P02768; 187H.
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.

```

```

DR PRINTS; P000802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON_TER 1 1
FT SIGNAL <1 16 BY SIMILARITY.
FT PROPEP 17 22 BY SIMILARITY.
FT CHAIN 23 605 SERUM ALBUMIN.
FT DOMAIN 23 202 ALBUMIN 1.
FT DOMAIN 209 394 ALBUMIN 2.
FT DOMAIN 401 592 ALBUMIN 3.
FT METAL 31 31 COPPER (BY SIMILARITY).
FT DISULFID 75 84 BY SIMILARITY.
FT DISULFID 97 113 BY SIMILARITY.
FT DISULFID 112 123 BY SIMILARITY.
FT DISULFID 145 190 BY SIMILARITY.
FT DISULFID 189 198 BY SIMILARITY.
FT DISULFID 221 267 BY SIMILARITY.
FT DISULFID 266 274 BY SIMILARITY.
FT DISULFID 286 300 BY SIMILARITY.
FT DISULFID 299 310 BY SIMILARITY.
FT DISULFID 337 382 BY SIMILARITY.
FT DISULFID 381 390 BY SIMILARITY.
FT DISULFID 413 459 BY SIMILARITY.
FT DISULFID 458 469 BY SIMILARITY.
FT DISULFID 482 498 BY SIMILARITY.
FT DISULFID 497 508 BY SIMILARITY.
FT DISULFID 535 580 BY SIMILARITY.
FT DISULFID 579 588 BY SIMILARITY.
FT CONFLICT 562 562 E->D (IN REF. 1; AAA30988).
SQ SEQUENCE 605 AA; 69410 MW; 3E55680DD1A1F4FF CRC64;

Query Match 100.0%; Score 38; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
DB 76 VADESAEN 83

RESULT 3
ALBU_HUMAN STANDARD; PRT; 609 AA.
AC P02768; O55574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9UJZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Rufner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczky A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
RC MEDLINE=82081982; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GUY-121.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 0.583564 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_54_61
Perfect score: 38
Sequence: 1 VADESAEN 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	38	100.0	600	1 ALBU MACMU
2	38	100.0	605	1 ALBU PIG
3	38	100.0	609	1 ALBU HUMAN
4	35	92.1	608	1 ALBU RAT
5	34	89.5	607	1 ALBU HORSE
6	32	84.2	608	1 ALBU MOUSE
7	32	84.2	608	1 ALBU RABIT
8	31	81.6	411	1 LAGI YEAST
9	31	81.6	442	1 JRKL HUMAN
10	30	78.9	478	1 SYH XANAC
11	30	78.9	609	1 ALBU MERUN
12	30	78.9	651	1 SECS YEAST
13	30	78.9	868	1 SRYC DROME
14	29	76.3	78	1 ACP EUCBP
15	29	76.3	114	1 PARA TRYBS
16	29	76.3	186	1 KPTA AGRT5
17	29	76.3	263	1 MAZG ECOLI
18	29	76.3	311	1 SREL CAEEL
19	29	76.3	409	1 GATD THEAC
20	29	76.3	410	1 CGEL HUMAN
21	29	76.3	425	1 HISX HALN1
22	29	76.3	490	1 AMPI STRCO
23	29	76.3	608	1 ALBU FELCA
24	29	76.3	1148	1 RFCL HUMAN
25	29	76.3	1406	1 TOPI CANGA
26	29	76.3	1463	1 PAZL BOVIN
27	29	76.3	2003	1 YDBA ECOLI
28	28	73.7	76	1 YNCU ECOLI
29	28	73.7	148	1 FLAV DESDE
30	28	73.7	158	1 PRI ASPOF
31	28	73.7	179	1 GRPE STRMU
32	28	73.7	212	1 SN25 DROME
33	28	73.7	280	1 AROK ARCFU

RESULT 1

ID	ALBU MACMU	STANDARD;	PRT;	600 AA.
AC	Q28522;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Serum albumin precursor (Fragment).			
GN	ALB.			
OS	Macaca mulatta (Rhesus macaque).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;			
OC	Cercopitheidae; Macaca.			
OX	NCBI_TaxID=9544;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93211971; PubMed=8460152;			
RA	Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,			
RA	Dwulet J., Putnam F.W.			
RT	"cDNA and protein sequence of polymorphic macaque albumins that differ			
RT	in bilirubin binding".			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).			
CC	!- FUNCTION: Serum albumin, the main protein of plasma, has a good			
CC	binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,			
CC	hormones, bilirubin and drugs. Its main function is the regulation			
CC	of the colloidal osmotic pressure of blood.			
CC	!- TISSUE SPECIFICITY: Plasma.			
CC	!- SUBCELLULAR LOCATION: Secreted.			
CC	!- SIMILARITY: Belongs to the ALB/APP/VDB family.			
CC	!- SIMILARITY: Contains 3 albumin domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@sb-sib.ch).			
DR	EMBL; M90463; AAA36906.1; -			
DR	PIR; A47391; A47391.			
DR	HSP; P02768; 1E7B.			
DR	InterPro; IPR000264; Serum_albumin.			
DR	Pfam; PF00273; transport_prot; 3.			
DR	PRINTS; P00802; SERUMALBUMIN.			
DR	ProDom; PD002486; Serum_albumin; 1.			
DR	SMART; SM00103; ALBUMIN; 3.			
DR	PROSITE; PS00212; ALBUMIN; 3.			
KW	Metal-binding; Lipid-binding; Repeat; Signal; Copper.			
FT	NON_TER 1			
SIGNAL	<1 10			
PROPEP	11 16			
FT	CHAIN 17 600			
FT	DOMAIN 17 197			
FT	DOMAIN 204 389			
FT	DOMAIN 396 587			

P29571 methanobact
Q889m3 pseudomonas
P52350 human herpe
P34392 caenorhabdi
Q830u2 enterococcu
O59636 pyrococcus
Q9ptn4 xenopus lae
P44001 haemophilus
P30183 arabidopsis
Q8p28 methanosarc
P05099 escherichia
Q92q12 rhizobium m

ALIGNMENTS

Best Local Similarity 87.5%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8
DB 176 VADSAEN 183

QY 2 ADESAEN 8
DB 375 ADESSSEN 381

Search completed: April 19, 2004, 12:02:22
Job time : 2.97507 secs

RESULT 14

S46800
LAG1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YHL003C
C:Species: Saccharomyces cerevisiae
C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 23-Mar-2001
C:Accession: S46800; A54012
R:Favella, T.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9780.
A:Reference number: S46797
A:Accession: S46800
A:Molecule type: DNA
A:Residues: 1-411 <FAV>
A:Cross-references: EMBL:U10555; NID:G500813; PIDN:AA68429.1; PID:G500820; MIPS:YHL003C
R:D'Amico, N.P.; Childress, A.M.; Franklin, D.S.; Kale, S.P.; Pinskiwasdi, C.; Jazwinski,
J. Biol. Chem. 269, 15451-15459, 1994
A:Title: Cloning and characterization of LAG1, a longevity-assurance gene in yeast.
A:Reference number: A54012; MUID:94253121; PMID:8195187
A:Accession: A54012
A:Molecule type: DNA
A:Residues: 1-172, 'IV', 175-219, 'C', 221-300, 'TEISGI', 314, 'EKQE', 315, 'DSNDNPTE', 324, 'A', 32
'A', 381, 'AGQR', 386, 'L', 388, 'NRLARNNEK' <DAM>
A:Cross-references: GB:U08133
C:Genetics:
A:Gene: SGD:LAG1
A:Cross-references: SGD:S0000995; MIPS:YHL003C
A:Map position: 8L
C:Function:
A:Description: involved in determination of longevity
C:Superfamily: hypothetical protein YXL008C
C:Keywords: transmembrane protein

Query Match 81.6%; Score 31; DB 2; Length 411;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8
DB 397 SDESAEN 403

RESULT 15

JCS594
jerky gene protein homolog - human
C:Species: Homo sapiens (man)
C>Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 05-Nov-1999
C:Accession: JCS594
R:Zeng, Z.; Kyaw, H.; Gakenheimer, K.R.; Augustus, M.; Fan, P.; Zhang, X.; Su, K.; Carter,
Biochem. Biophys. Res. Commun. 236, 389-395, 1997
A:Title: Cloning, mapping, and tissue distribution of a human homologue of the mouse jer
A:Reference number: JCS594; MUID:97382443; PMID:9240447
A:Accession: JCS594
A:Molecule type: mRNA
A:Residues: 1-442 <ZEN>
A:Cross-references: DBJ:AF004715; NID:G2314828; PIDN:AA65833.1; PID:G2314829
A>Note: it is uncertain whether Met-1 or Met-33 is the initiator
C:Comment: This protein functions as a nuclear regulatory protein.
C:Genetics:
A:Gene: hhmig
A:Map position: 11q21

Query Match 81.6%; Score 31; DB 2; Length 442;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 231 VDSQAQN 238

RESULT 9

T21861

hypothetical protein F36P2.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21861

R:Cottage, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19479

A:Accession: T21861

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1877 <WIL>

A:Cross-references: EMBL:Z81532; PIDN:CAB04326.1; GSPDB:GN00019; CESP:F36P2.3

A:Experimental source: clone F36P2

C:Genetics:

A:Gene: CESP:F36P2.3

A:Map position: 1

A:Introns: 49/3; 86/2; 112/3; 139/2; 235/3; 284/3; 436/3; 507/2; 566/2; 678/2; 1000/3; 1000/3

Query Match 84.2%; Score 32; DB 2; Length 1877;

Best Local Similarity 75.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8

Db 714 VADENAQN 721

RESULT 10

T16871

hypothetical protein T13H2.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000

C:Accession: T16871

R:Wu, X.

submitted to the EMBL Data Library, October 1995

A:Description: The sequence of C. elegans cosmid T13H2.

A:Reference number: Z18593

A:Accession: T16871

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2215 <WUX>

A:Cross-references: EMBL:U39653; NID:G1049397; PID:G1049401; PIDN:AB52495.1; GSPDB:GN000

A:Experimental source: strain Bristol N2; clone T13H2

C:Genetics:

A:Gene: CESP:T13H2.4

A:Map position: X

A:Introns: 112/2; 136/1; 167/1; 196/1; 649/1; 696/1; 757/1; 850/1; 882/1; 985/1; 1046/1;

Query Match 84.2%; Score 32; DB 2; Length 2215;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAAE 7

Db 129 VADESAAE 135

RESULT 11

AC1251

hypothetical protein lmo1411 [imported] - Listeria monocytogenes (strain EGD-e)

C:Species: Listeria monocytogenes

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AC1251

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.N.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1251

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAC99489.1; PID:gl6410840; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo1411

Query Match 81.6%; Score 31; DB 2; Length 75;

Best Local Similarity 75.0%; Pred. No. 8.8;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8

Db 66 VADDSVEN 73

RESULT 12

D71442

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

A:Variety: columbia

C>Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998

C:Accession: D71442

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dir

F.; Wedler, H.; Wedler, E.; Wambutt, R.; Weizensegger, T.; Pohl, T.M.; Terry, N.; Gie

avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenec

erhoft, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ar

C.; Chludzisz, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis th

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: D71442

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-232 <BEV>

A:Cross-references: GB:Z97343; NID:G2245073; PID:e327054; PID:G2245091

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match 81.6%; Score 31; DB 2; Length 232;

Best Local Similarity 85.7%; Pred. No. 29;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8

Db 117 SDESAEN 123

RESULT 13

T36578

probable membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T36578

R:Oliver, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z21575

A:Accession: T36578

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-282 <OLI>

A:Cross-references: EMBL:AL049826; PIDN:CAB42715.1; GSPDB:GN00070; SCOREDB:SCH24.13C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOREDB:SCH24.13C

Query Match 81.6%; Score 31; DB 2; Length 282;

Mol. Cell. Biol. 7, 2425-2434, 1987
A>Title: Determinants of rat albumin promoter tissue specificity analyzed by an improved
A:Reference number: 157621; MUID:87286876; PMID:3475566
A:Accession: 157621
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5 <RES>
A:Cross-references: GB:M16825; NID:g202828; PIDN:AAA0712.1; PID:g554412
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status experimental <SIG>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-608/Product: serum albumin #status experimental
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status experimental
F:77-86,99-115,114-125,148-153,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4
F:77-86,99-115,114-125,148-153,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4

Query Match 92.1%; Score 35; DB 1; Length 608;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 78 VADESAEN 85
|||||

RESULT 5
ABHOS
serum albumin precursor - horse
C:Species: Equus caballus (domestic horse)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S34053
R:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
Eur. J. Biochem. 215, 205-212, 1993
A>Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
A:Reference number: S34053; MUID:93345495; PMID:8344282
A:Accession: S34053
A:Molecule type: mRNA
A:Residues: 1-607 <HOA>
A:Cross-references: GB:X74045; NID:g399671; PIDN:CAA52194.1; PID:g399672
C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4

Query Match 89.5%; Score 34; DB 1; Length 607;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADESAEN 8
Db 79 ADESAEN 85
|||||

RESULT 6
T44794
hypothetical protein [imported] - Haloflex alicantei
C:Species: Haloflex alicantei
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T44794
R:Holmes, M.L.; Dyall-Smith, M.L.
submitted to the EMBL Data Library, May 1999

A:Reference number: Z22843
A:Accession: T44794
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-157 <HOL>
A:Cross-references: EMBL:U70664; PIDN:AAB40124.1
A:Experimental source: strain SBI

Query Match 84.2%; Score 32; DB 2; Length 157;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 102 VADESAEN 109
|||||

RESULT 7
G81713
RNA polymerase sigma factor, sigma-70 family TC0331 [imported] - Chlamydia muridarum (s
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: G81713
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A>Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: G81713
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-253 <TET>
A:Cross-references: GB:AE002301; GB:AE002160; NID:g7190372; PIDN:AAF39194.1; PID:g71903
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0331
C:Superfamily: transcription initiation factor sigmaD; transcription initiation factor

Query Match 84.2%; Score 32; DB 2; Length 253;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 174 IADERAEN 181
|||||

RESULT 8
T45966
hypothetical protein F7J8.240 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45966
R:Bevan, M.; Zimmermann, W.; Gruenisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Le
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23018
A:Accession: T45966
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-385 <BEV>
A:Cross-references: EMBL:AL137189
A:Experimental source: cultivar Columbia; BAC clone F708
C:Genetics:
A:Map position: 5
A:Introns: 72/3; 136/3; 303/3
A>Note: F7J8.240

Query Match 84.2%; Score 32; DB 2; Length 385;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
|||||

A;Molecule type: protein
A;Residues: 82-105, 'K', 107-110 <GAL2>
A;Note: this variant is designated albumin Vibo Valentia
A;Accession: A38255
A;Molecule type: protein
A;Residues: 76-83, 'K', 85-106 <GAL3>
A;Note: this variant is designated albumin Torino
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin Her
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: protein
A;Residues: 255-263, 'E', 265-281 <MIN1>
A;Note: this variant is designated albumin Herborn
R;Minchiotti, L.; Galliano, M.; Scoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta,
Biochim. Biophys. Acta 1119, 232-238, 1992
A;Title: Two albumins with identical electrophoretic mobility are produced by differ
A;Reference number: S21078; MUID:92190239; PMID:1347703
A;Accession: S21078
A;Molecule type: protein
A;Residues: 354-356, 'K', 358-378 <MIN2>
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported,
R;He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A;Title: Atomic structure and chemistry of human serum albumin.
A;Reference number: A46756; MUID:92334427; PMID:1630489
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.O.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40,
A;Reference number: A94442
A;Contents: annotation; Three-dimensional structure and disulfide bonds
R;Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A;Title: Disulfide bonds in human serum albumin.
A;Reference number: A90930
A;Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A;Reference number: A90299; MUID:78186630; PMID:656055
A;Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjöholm, I., eds., 11-20,
A;Title: Serum albumin: conformation and active sites.
A;Reference number: A94408
A;Contents: annotation; binding sites
R;Harper, M.E.; Dugaiczky, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes
A;Reference number: A90028; MUID:83279982; PMID:6192711
A;Contents: annotation; gene position
R;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.
A;Reference number: A46755; MUID:76257808; PMID:955075
A;Contents: annotation
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic acid
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A;Title: Identification of Lys(199) as the primary binding site for pyridoxal 5'-phospha
A;Reference number: A56294; MUID:92183881; PMID:1544460
A;Contents: annotation
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p
tase activity
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
C;Comment: A large number of variants of human serum albumin have been described.
C;Genetics:
A;Gene: GDB:ALB
A;Cross-references: GDB:118990; OMIM:103600
A;Map position: 4q11-4q13
C;Superfamily: serum albumin; serum albumin repeat homology

C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyrid
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-609/Product: serum albumin #status experimental <MPT>
F;229-202/Domain: serum albumin repeat homology <SA1>
F;166-174/Product: Kinetensin #status experimental <KIP>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,
F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental
Query Match 100.0%; Score 38; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. NO. 2.2; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0;
QY 1 VADESAEN 8
DB 78 VADESAEN 85
RESULT 4
ABRTS
serum albumin precursor - rat
N;Alternate names: preproalbumin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999
C;Accession: A93872; A92211; A91940; C45800; I57621; A03233
R;Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A;Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A;Reference number: A93872; MUID:81223722; PMID:7017712
A;Accession: A93872
A;Molecule type: mRNA
A;Residues: 1-608 <SAR>
A;Cross-references: GB:V01222; GB:J00698; NID:955627; PID:CAA24532.1; PID:955628
R;Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A;Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Anal
A;Reference number: A92211; MUID:77249557; PMID:893447
A;Note: cleavages during protein maturation
A;Accession: A92211
A;Molecule type: protein
A;Residues: 1-38 <STR>
R;Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A;Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleave
A;Reference number: A91946; MUID:78109425; PMID:564345
A;Accession: A91946
A;Molecule type: protein
A;Residues: 223-288; 572-608 <IS2>
A;Note: 262-Leu was also found
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.
A;Reference number: A90758; MUID:79001617; PMID:80265
A;Contents: annotation; copper binding
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid protea
A;Reference number: A45800; MUID:89341405; PMID:2474609
A;Accession: C45800
A;Status: preliminary
A;Molecule type: protein
A;Residues: 166-173 <CAR>
R;Heard, J.

A:Molecule type: protein
A:Residues: 23-51,'X','53-54','XXXY' 146,'E',149,'E',150-151,'XVY',155 <LIM>
A:Experimental source: dental enamel
A>Note: Albumin and other serum proteins are also found in bone
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
F:17-22/Domain: propeptide #status predicted <PRO>
F:23-605/Product: serum albumin #status predicted <MAT>
F:27-199/Domain: serum albumin repeat homology <SA1>
F:218-391/Domain: serum albumin repeat homology <SA2>
F:410-589/Domain: serum albumin repeat homology <SA3>
F:75-84,97-113,112-123,145-190,189-198,221-267,266-274,286-300,299-310,337-382,381-390,4
F:261/Binding site: bilirubin (Lys) #status predicted

Query Match 100.0%; Score 38; DB 1; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
|||||||
Db 76 VADESAEN 83

RESULT 3
ABHUS
serum albumin precursor [validated] - human
N:Alternate names: preproalbumin
N:Contains: kinetensin
C:Species: Homo sapiens (man)
C>Date: 29-Jul-1991 #sequence, revision 31-Jan-1997 #text change 17-Mar-2000
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S36
R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebur
Nucleic Acids Res. 9, 6103-6114, 1981
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia coli
A:Reference number: A93743; MUID:82081882; PMID:6171778
A:Accession: A93743
A:Molecule type: mRNA
A:Residues: 1-419,'K',421-609 <LAW>
A:Cross-references: EMBL:V00495; GB:J000078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA23
R:Dugaiczky, A.; Law, S.W.; Dennis, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A:Reference number: A93936; MUID:82105994; PMID:6275391
A:Accession: A93936
A:Molecule type: mRNA
A:Residues: 1-120,'G',122-609 <DUG>
A:Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590
R:Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and
A:Reference number: I39427; MUID:86140099; PMID:2419329
A:Accession: I39427
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-26 <URA>
A:Cross-references: GB:M13075; NID:G178330; PIDN:AAAS1688.1; PID:G553173
R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family.
A:Reference number: I59286; MUID:94181575; PMID:8134387
A:Accession: I59286
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 282-290,'KSRPDLQ' <WAT>
A:Cross-references: GB:G59152; NID:G546032; PIDN:AA30282.1; PID:G546033
R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl-
A:Reference number: I59313; MUID:94294404; PMID:8022807

A:Accession: I59313
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 589-590,'ALPRRVKNLLQVKLP' <WAD>
A:Cross-references: GB:S70799; NID:G547231; PIDN:AA31177.1; PID:G547232
A>Note: This frame-shift variant, designated albumin Bazzano, four additional variant
R:Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A:Reference number: G08292
A:Accession: G01747
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-120,'G',122-455 <MEN>
A:Cross-references: EMBL:U22961; NID:G763428; PIDN:AAA64922.1; PID:G763431
R:Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A:Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Ke
A:Reference number: S55314; MUID:95275251; PMID:7755581
A:Accession: S55314
A:Molecule type: protein
A:Residues: 19-27 <LED>
R:Meloun, B.; Moravsek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A:Title: Complete amino acid sequence of human serum albumin.
A:Reference number: A91420; MUID:76187907; PMID:1225573
A:Accession: A91420
A:Molecule type: protein
A:Residues: 25-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',399;
R:Roehr, U.; Spitteller, G.; Tripier, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A:Title: Isolation and structure elucidation of middle-molecular weight peptides from t
A:Reference number: S06422
A>Note: this paper is in German, with an English abstract
A:Accession: S06422
A:Molecule type: protein
A:Residues: 25-48 <ROE>
R:Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A:Title: Mass spectrometric identification of modifications to human serum albumin tre
A:Reference number: S36882; MUID:93384321; PMID:8373198
A:Accession: S36882
A:Molecule type: protein
A:Residues: 45-67,141-160;311-337,469-490;570-581 <FIN>
R:Kausler, E.; Spitteller, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A:Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelme
A:Reference number: S17599; MUID:92126241; PMID:1772598
A:Accession: S17599
A:Molecule type: protein
A:Residues: 25-54;354-357;431-447 <KAU>
A>Note: 49-Leu was also found
R:Carraway, R.E.; Cochrane, D.E.; Boucher, W.; Mittra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid protea
A:Reference number: A45800; MUID:89341406; PMID:2474609
A:Accession: A45800
A:Molecule type: protein
A:Residues: 166-173 <CAR>
R:Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Y
Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A:Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-t
A:Reference number: A03239; MUID:86242180; PMID:3087352
A:Accession: A03239
A:Molecule type: protein
A:Residues: 166-173,'L' <MOG>
R:Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins,
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A:Title: Mutations in genetic variants of human serum albumin found in Italy.
A:Reference number: A38255; MUID:91062352; PMID:2247440
A:Accession: C38255
A:Molecule type: protein
A:Residues: 76-111 <GAL1>
A:Accession: B38255

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 0.975069 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_54_61
Perfect score: 38
Sequence: 1 VADESAEN 8
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:.*
1: Pirl:.*
2: Pirl2:.*
3: Pirl3:.*
4: Pirl4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	600	2 A47391	serum albumin prec
2	38	100.0	603	1 ABPGS	serum albumin prec
3	38	100.0	609	1 ABHUS	serum albumin prec
4	35	92.1	608	1 ABRTS	serum albumin prec
5	34	89.5	607	1 ABHOS	serum albumin prec
6	32	84.2	157	2 T44794	hypothetical prote
7	32	84.2	253	2 G81713	RNA polymerase sig
8	32	84.2	385	2 T45966	hypothetical prote
9	32	84.2	1877	2 T21861	hypothetical prote
10	32	84.2	2215	2 T16871	hypothetical prote
11	31	81.6	75	2 AC1251	hypothetical prote
12	31	81.6	232	2 D71442	hypothetical prote
13	31	81.6	282	2 T36578	probable membrane
14	31	81.6	411	2 S46800	LAG1 protein - yea
15	31	81.6	442	2 JC5594	jerky gene protein
16	31	81.6	785	2 T00474	hypothetical prote
17	31	81.6	1657	2 T25421	hypothetical prote
18	30	78.9	109	2 S32886	hypothetical prote
19	30	78.9	221	2 C94899	hypothetical prote
20	30	78.9	265	2 D84138	endonuclease III
21	30	78.9	309	2 P86393	protein T24P13.4
22	30	78.9	387	2 D84885	hypothetical prote
23	30	78.9	465	2 T22141	hypothetical prote
24	30	78.9	475	2 D84533	hypothetical prote
25	30	78.9	496	2 AD1211	B. subtilis Yuek p
26	30	78.9	496	2 AD1567	conserved hypothet
27	30	78.9	508	2 T22954	hypothetical prote
28	30	78.9	609	2 JC5838	albumin - Mongolia
29	30	78.9	651	2 A55100	SEC9 protein - yea

serendipity (sry h
procyelic acidic r
conserved hypothet
probable RNA 2'-ph
transcription regu
mazG protein - Esc
hypothetical prote
hypothetical prote
protein B0495.1 [i
probable oxygen-in
cyclin E - human
histidinol dehydro
probable membrane
hypothetical prote
X-Pro aminopeptida
LlDBP protein - hu

ALIGNMENTS

RESULT 1

A47391
serum albumin precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A47391
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwyer, J.; Putnam, I
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bili:
A:Reference number: A47391; MUID:93211971; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-600 <WAT>
A:Cross-references: GB:M30463; NID:G342294; PIDN:AAA36906.1; PID:G342295
A:Experimental source: liver
A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIP:128281)
C:Superfamily: serum albumin; serum albumin repeat homology
F:21-194/Domain: serum albumin repeat homology <SA1>
F:213-386/Domain: serum albumin repeat homology <SA2>
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 38; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
|||
Db 70 VADESAEN 77

RESULT 2

ABPGS
serum albumin precursor - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S01382; A61006
R:Reinstock, J.; Baldwin, G.S.
Nucleic Acids Res 16, 9045, 1988
A:Title: Nucleotide sequence of porcine liver albumin.
A:Reference number: S01382; MUID:89016582; PMID:3174440
A:Accession: S01382
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-605 <WEI>
A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798
R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
J. Bone Miner. Res. 4, 235-241, 1989
A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of miner:
A:Reference number: A61006; MUID:89269769; PMID:2728927
A:Accession: A61006

Db 307 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLECCCKPLEKSHCIAEVENDEMPA 366
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLEKC 360
Db 367 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLEKC 426
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387
Db 427 CAAADPHECYAKVDFEFKPLVEEPQNL 453

RESULT 15

US-10-153-064-105
; Sequence 105, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 105
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-105

Query Match 100.0%; Score 2068; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 2.1e-203; Indels 0; Gaps 0;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHKSEVAHRFKDIDGENFKALVLIAPAQYLQCCPEDHVKLVNEVTEFAKTCVADESAE 60
Db 67 DAHKSEVAHRFKDIDGENFKALVLIAPAQYLQCCPEDHVKLVNEVTEFAKTCVADESAE 126
Qy 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
Db 127 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 186
Qy 121 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQADKAAACLLP 180
Db 187 DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQADKAAACLLP 246
Qy 181 KLDELDRDEGKASSAKQBLKCSLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
Db 247 KLDELDRDEGKASSAKQBLKCSLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 306
Qy 241 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLECCCKPLEKSHCIAEVENDEMPA 300
Db 307 VHTCCGGDLLECCADDDRADLAKYICENQDSISSKLECCCKPLEKSHCIAEVENDEMPA 366
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLEKC 360
Db 367 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLEKC 426
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387
Db 427 CAAADPHECYAKVDFEFKPLVEEPQNL 453

Search completed: April 19, 2004, 12:05:18
Job time : 60.3186 secs

```
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-153-064-133

Query Match      100.0%; Score 2068; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 2e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 67 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 126
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 186
QY 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 187 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 246
QY 181 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFPKAFSAEVSCLVTDLT 240
DB 247 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFPKAFSAEVSCLVTDLT 306
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
DB 307 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 366
QY 301 DLPSLAADFVESKDVCKNYAAEKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETTLEKC 360
DB 367 DLPSLAADFVESKDVCKNYAAEKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETTLEKC 426
QY 361 CAAADPHECVAKVFDEPKPLVEBPQNL 387
DB 427 CAAADPHECVAKVFDEPKPLVEBPQNL 453

RESULT 14
US-10-153-064-99
; Sequence 99, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 99
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-153-064-99

Query Match      100.0%; Score 2068; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 2.1e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 67 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 126
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 186
QY 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 187 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 246
QY 181 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFPKAFSAEVSCLVTDLT 240
DB 247 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFPKAFSAEVSCLVTDLT 306
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

Query Match      100.0%; Score 2068; DB 4; Length 652;
Best Local Similarity 100.0%; Pred. No. 2.1e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 67 DAHKSEVAHRFKDLGEENFKALVLIAPAQYLOQCPFEDHVKLVNEVTEFAKTCVADESAAE 126
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 127 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 186
QY 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 187 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 246
QY 181 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFPKAFSAEVSCLVTDLT 240
DB 247 KLDELDRDEGKASSAKQRLKCSAQKGERAFKAWARLSQRPFPKAFSAEVSCLVTDLT 306
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
```

PRIOR APPLICATION DATA: US 08/256,927
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-2

Query Match 100.0%; Score 2068; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAE 84
Qy 61 NCDSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 204
Qy 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
Db 205 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264
Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
Db 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 324
Qy 301 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLLAKTYETTLK 360
Db 325 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLLAKTYETTLK 384
Qy 361 CAAADPHECYAKVDFEFPKPLVEEPQNL 387
Db 385 CAAADPHECYAKVDFEFPKPLVEEPQNL 411

PRIOR APPLICATION DATA: US 08/256,927
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-797-689-2

Query Match 100.0%; Score 2068; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAE 60
Db 25 DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNEVTEFAKTCVADESAE 84
Qy 61 NCDSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMTAFHDNEETFLKKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 204
Qy 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
Db 205 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 264
Qy 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
Db 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 324
Qy 301 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLLAKTYETTLK 360
Db 325 DLPSLAADFVSKDKVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLLAKTYETTLK 384
Qy 361 CAAADPHECYAKVDFEFPKPLVEEPQNL 387
Db 385 CAAADPHECYAKVDFEFPKPLVEEPQNL 411

RESULT 11
US-09-984-186-2
Sequence 2, Application US/09984186
Patent No. 6686179
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guillon, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43

RESULT 12
US-10-153-064-133

```

; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
; US-09-976-594-977

Query Match 100.0%; Score 2068; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDGLGEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRPKDGLGEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 144
QY 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 264
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECKECPLEKSHCIAEVENDEMPA 300
DB 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECKECPLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 384
QY 361 CAADPHECYAKVDFEFPKLVPEPQNL 387
DB 385 CAADPHECYAKVDFEFPKLVPEPQNL 411

RESULT 9
PCT-US95-04075-3
; Sequence 3, Application PC/TUS9504075
; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRR
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04075
; FILING DATE:

```

```

; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US95-04075-3

Query Match 100.0%; Score 2068; DB 5; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDGLGEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHSEVAHRPKDGLGEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 144
QY 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 205 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 264
QY 241 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECKECPLEKSHCIAEVENDEMPA 300
DB 265 VHTCCHGDLLECADRADLAKYICENQDSISSKLKECKECPLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 384
QY 361 CAADPHECYAKVDFEFPKLVPEPQNL 387
DB 385 CAADPHECYAKVDFEFPKLVPEPQNL 411

RESULT 10
US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Pournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435

```

Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEPAEVSCLVTDLTJK 240
Db 205 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEPAEVSCLVTDLTJK 264
Qy 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 360
Db 325 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 384
Qy 361 CAADAPHECYAKVDFEKPFLVEEPQNL 387
Db 385 CAADAPHECYAKVDFEKPFLVEEPQNL 411

RESULT 6

US-08-897-956A-2

; Sequence 2, Application US/08897956A

; Patent No. 6423512

; GENERAL INFORMATION:

; APPLICANT: Mary Ellen Digan

; APPLICANT: Philip Lake

; APPLICANT: Hermann Gram

; TITLE OF INVENTION: Fusion Polypeptides

; FILE REFERENCE: 600-7244/CPA

; CURRENT APPLICATION NUMBER: US/08/897,956A

; CURRENT FILING DATE: 1997-07-21

; PRIOR APPLICATION NUMBER: 60/022,689

; PRIOR FILING DATE: 1996-07-26

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 609

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-08-897-956A-2

Query Match 100.0%; Score 2068; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEPAEVSCLVTDLTJK 240
Db 205 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEPAEVSCLVTDLTJK 264

Qy 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 360
Db 325 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 384
Qy 361 CAADAPHECYAKVDFEKPFLVEEPQNL 387
Db 385 CAADAPHECYAKVDFEKPFLVEEPQNL 411

RESULT 7

US-10-153-064-7

; Sequence 7, Application US/10153064

; Patent No. 6663485

; GENERAL INFORMATION:

; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064

; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 7

; LENGTH: 609

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-153-064-7

Query Match 100.0%; Score 2068; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
Qy 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLP 204
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEPAEVSCLVTDLTJK 240
Db 205 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORFPKAEPAEVSCLVTDLTJK 264
Qy 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
Db 265 VHTCCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 324
Qy 301 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 360
Db 325 DLPSLAADPFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 384
Qy 361 CAADAPHECYAKVDFEKPFLVEEPQNL 387
Db 385 CAADAPHECYAKVDFEKPFLVEEPQNL 411

RESULT 8

US-09-976-594-977

; Sequence 977, Application US/09976594

; Patent No. 6673549

; GENERAL INFORMATION:

; APPLICANT: Furness, Michael

; APPLICANT: Buchbinder, Jenny

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MERRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-769-746-2

Query Match      100.0%; Score 2068; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.8e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
|
|
|
Db 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
|
|
|
QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHQDDNPNLPRLVRPEV 120
|
|
|
Db 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHQDDNPNLPRLVRPEV 120
|
|
|
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
|
|
|
Db 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
|
|
|
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPFKAFAEVSKLVTDLTK 240
|
|
|
Db 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPFKAFAEVSKLVTDLTK 240
|
|
|
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKEPLLEKSHGICIAEVNDEMPPA 300
|
|
|
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKEPLLEKSHGICIAEVNDEMPPA 300
|
|
|
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLAKTYTTLEK 360
|
|
|
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLAKTYTTLEK 360
|
|
|
QY 361 CAAADPHECYAKVFDEKPLVEBPQNL 387
|
|
|
Db 361 CAAADPHECYAKVFDEKPLVEBPQNL 387
|
|
|

RESULT 5
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESS: Amgen Center, Patent Operations/RR
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-222-619-3

Query Match      100.0%; Score 2068; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.9e-203;

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10153064
; FILING DATE: 2002-05-24
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bell et al.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MERRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-769-746-2

Query Match      100.0%; Score 2068; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.8e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
|
|
|
Db 1 DAHKSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPPEHDHVKLVNEVTEFAKTCVADESAAE 60
|
|
|
QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHQDDNPNLPRLVRPEV 120
|
|
|
Db 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFLOHQDDNPNLPRLVRPEV 120
|
|
|
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
|
|
|
Db 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
|
|
|
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPFKAFAEVSKLVTDLTK 240
|
|
|
Db 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPFKAFAEVSKLVTDLTK 240
|
|
|
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKEPLLEKSHGICIAEVNDEMPPA 300
|
|
|
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCECKEPLLEKSHGICIAEVNDEMPPA 300
|
|
|
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLAKTYTTLEK 360
|
|
|
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVLRLAKTYTTLEK 360
|
|
|
QY 361 CAAADPHECYAKVFDEKPLVEBPQNL 387
|
|
|
Db 361 CAAADPHECYAKVFDEKPLVEBPQNL 387
|
|
|

RESULT 4
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-153-064-5
```

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Region
; LOCATION: 389..419
; OTHER INFORMATION: /note= "Alternative C-termini of
; OTHER INFORMATION: HSA(1-n)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..585
; OTHER INFORMATION: /note= "Amino acid sequence of
; OTHER INFORMATION: natural HSA"
US-08-153-799-14

Query Match 100.0%; Score 2068; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.8e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELREGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPPKAEFAEVSKLVTDLTK 240
DB 181 KLDELREGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCGGDILLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGGDILLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360

RESULT 2
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19408-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match 100.0%; Score 2068; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.8e-203;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGEENFKALVLIAPQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELREGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPPKAEFAEVSKLVTDLTK 240
DB 181 KLDELREGKASSAKQRLKCSLQKGERAFKAWAVARLSQRPPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCGGDILLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGGDILLECADRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360

RESULT 3
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 59.3186 Seconds
(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_1_387
Perfect score: 2068
Sequence: 1 DAHKSEVAHRFKDLGEENPK.....ECYAKVFDEKPLVEEPQNL 387

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/backfile1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2068	100.0	585	1	US-08-153-799-14
2	2068	100.0	585	2	US-08-702-572-2
3	2068	100.0	585	3	US-08-769-746-2
4	2068	100.0	585	4	US-10-153-064-5
5	2068	100.0	609	1	US-08-222-619-3
6	2068	100.0	609	4	US-08-897-956A-2
7	2068	100.0	609	4	US-10-153-064-7
8	2068	100.0	609	4	US-09-976-594-977
9	2068	100.0	609	5	PCT-US95-04075-3
10	2068	100.0	610	2	US-08-797-689-2
11	2068	100.0	610	4	US-08-984-186-2
12	2068	100.0	651	4	US-10-153-064-133
13	2068	100.0	652	4	US-10-153-064-96
14	2068	100.0	652	4	US-10-153-064-99
15	2068	100.0	652	4	US-10-153-064-105
16	2068	100.0	652	4	US-10-153-064-132
17	2068	100.0	653	4	US-10-153-064-131
18	2068	100.0	655	4	US-10-153-064-130
19	2068	100.0	660	4	US-10-153-064-90
20	2068	100.0	660	4	US-10-153-064-93
21	2068	100.0	668	4	US-10-153-064-102
22	2068	100.0	676	4	US-10-153-064-95
23	2068	100.0	676	4	US-10-153-064-98
24	2068	100.0	676	4	US-10-153-064-104
25	2068	100.0	676	4	US-10-153-064-127
26	2068	100.0	676	4	US-10-153-064-129
27	2068	100.0	677	4	US-10-153-064-125

Sequence 123, Appl
Sequence 92, Appl
Sequence 101, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 16, Appl
Sequence 16, Appl
Sequence 3, Appl
Sequence 89, Appl
Sequence 4, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 5, Appl
Sequence 4, Appl
Sequence 6, Appl
Sequence 7, Appl
Sequence 1, Appl
Sequence 8, Appl

28 2068 100.0 680 4 US-10-153-064-123
29 2068 100.0 684 4 US-10-153-064-92
30 2068 100.0 682 4 US-10-153-064-101
31 2068 100.0 783 1 US-08-256-938-2
32 2068 100.0 787 1 US-08-256-938-4
33 2068 100.0 787 2 US-08-797-689-16
34 2068 100.0 787 4 US-09-984-186-16
35 2068 100.0 978 4 US-08-897-956A-3
36 2068 100.0 1104 4 US-10-153-064-89
37 2068 99.8 609 1 US-08-433-037-4
38 2062 99.7 585 1 US-08-448-196A-3
39 2062 99.7 585 2 US-08-984-176-1
40 1658.5 80.2 583 1 US-08-448-196A-5
41 1622.5 78.5 583 1 US-08-448-196A-4
42 1613.5 78.0 583 1 US-08-448-196A-6
43 1610 77.9 584 1 US-08-448-196A-7
44 1568 75.8 582 1 US-08-134-638-1
45 726.5 35.1 579 1 US-08-448-196A-8

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 70; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TVATLRETYGEMAD 14
Db      76 TVATLRETYGEMAD 89

RESULT 15
US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34

Query Match      100.0%; Score 70; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 TVATLRETYGEMAD 14
Db      76 TVATLRETYGEMAD 89

Search completed: April 19, 2004, 12:54:58
Job time : 5.89474 secs
```

; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match 100.0%; Score 70; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 76 TVATLRETYGEMAD 89

RESULT 11

US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match 100.0%; Score 70; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 76 TVATLRETYGEMAD 89

RESULT 12

US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12

; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match 100.0%; Score 70; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 76 TVATLRETYGEMAD 89

RESULT 13

US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match 100.0%; Score 70; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 76 TVATLRETYGEMAD 89

RESULT 14

US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbitt, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105

```

1 STATE: DC
2 COUNTRY: USA
3 ZIP: 20005-3315
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: Floppy disk
6 COMPUTER: IBM PC compatible
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8 SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
9 CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: US/09/984,010
11 FILING DATE: 21-May-2002
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: US 09/091,873
14 FILING DATE: 25-JUN-1998
15 APPLICATION NUMBER: PCT/GB96/03164
16 FILING DATE: 19-DEC-1996
17 INFORMATION FOR SEQ ID NO: 26:
18 SEQUENCE CHARACTERISTICS:
19 LENGTH: 585 amino acids
20 TYPE: amino acid
21 STRANDEDNESS: <Unknown>
22 TOPOLOGY: linear
23 MOLECULE TYPE: protein
24 HYPOTHETICAL: NO
25 ANTI-SENSE: NO
26 SEQUENCE DESCRIPTION: SEQ ID NO: 26:
27 US-09-984-010-26

```

```
Query Match      100.0%; Score 70; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	TVATLRETYGEMAD	14
Db	76	TVATLRETYGEMAD	89

```

RESULT 7
US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US2003012547A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICATT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF345
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-041-18

```

```
Query Match      100.0%; Score 70; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	TVATLRETYGEMAD	14
Db	76	TVATLRETYGEMAD	89

RESULT 8
US-09-833-117-18

```

? Sequence 18, Application US/09833117
? Publication No. US20030171267A1
? GENERAL INFORMATION:
?
? APPLICANT: Rosen, Craig A.
? APPLICANT: Sadeghi, Homa
? APPLICANT: Prior, Christopher P.
? APPLICANT: Turner, Andrew J.
?
? TITLE OF INVENTION: Albumin Fusion Proteins
?
? FILE REFERENCE: PF543
? CURRENT APPLICATION NUMBER: US/09/833,117
? CURRENT FILING DATE: 2001-04-12
? PRIOR APPLICATION NUMBER: 60/229,358
? PRIOR FILING DATE: 2000-04-12
? PRIOR APPLICATION NUMBER: 60/256,931
? PRIOR FILING DATE: 2000-12-21
? PRIOR APPLICATION NUMBER: 60/199,384
? PRIOR FILING DATE: 2000-04-25
?
? NUMBER OF SEQ ID NOS: 36
?
? SOFTWARE: PatentIn Ver. 2.1
?
? SEQ ID NO 18
?
? LENGTH: 585
?
? TYPE: PRT
?
? ORGANISM: Homo Sapiens
?
? US-09-833-117-18

```

```
Query Match          100.0%; Score 70; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	TVATLRETYGEMAD	14
D _b	76	TVATLRETYGEMAD	89

```

RESULT 9
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyvax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: Dyx-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-322-445

```

```
Query Match      100.0%; Score 70; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

RESULT 10
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030190043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: sleep, Darrell
; APPLICANT: Turner, Andrew J.

```
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-27

Query Match          100.0%; Score 70; DB 13; Length 241;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 100 TVATLRETYGEMAD 113

RESULT 3
US-10-074-956-28
; Sequence 28, Application US/10074956
; Publication No. US2002019332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08193-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-28

Query Match          100.0%; Score 70; DB 13; Length 268;
Best Local Similarity 100.0%; Pred. No. 7.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 100 TVATLRETYGEMAD 113

RESULT 4
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
```

```
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match          100.0%; Score 70; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

RESULT 5
US-09-932-613-445
; Sequence 445, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT: DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445

Query Match          100.0%; Score 70; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

RESULT 6
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 5.89474 Seconds
(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_76_89

Perfect score: 70

Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	195	13	US-10-074-956-24
2	70	100.0	241	13	Sequence 24, Appl
3	70	100.0	268	13	Sequence 27, Appl
4	70	100.0	585	9	US-10-074-956-28
5	70	100.0	585	10	Sequence 28, Appl
6	70	100.0	585	10	Sequence 2, Appl
7	70	100.0	585	10	Sequence 445, App
8	70	100.0	585	10	Sequence 18, Appl
9	70	100.0	585	10	Sequence 18, Appl
10	70	100.0	585	10	Sequence 18, Appl
11	70	100.0	585	11	US-09-832-501-18
12	70	100.0	585	11	Sequence 18, Appl
13	70	100.0	585	12	US-10-424-999-11
14	70	100.0	585	12	Sequence 11, Appl
15	70	100.0	585	12	Sequence 31, Appl
					Sequence 34, Appl

```

16 70 100.0 585 13 US-10-153-064-5 Sequence 5, Appli
17 70 100.0 585 14 US-10-153-604A-5 Sequence 5, Appli
18 70 100.0 585 14 US-10-319-263-1 Sequence 1, Appli
19 70 100.0 585 14 US-10-319-263-2 Sequence 2, Appli
20 70 100.0 585 14 US-10-414-469-1 Sequence 1, Appli
21 70 100.0 585 14 US-10-414-469-2 Sequence 2, Appli
22 70 100.0 585 14 US-10-413-831-1 Sequence 1, Appli
23 70 100.0 585 14 US-10-413-831-2 Sequence 2, Appli
24 70 100.0 585 15 US-10-413-832-1 Sequence 1, Appli
25 70 100.0 585 15 US-10-413-832-2 Sequence 2, Appli
26 70 100.0 585 15 US-10-414-386-1 Sequence 1, Appli
27 70 100.0 585 15 US-10-414-386-2 Sequence 2, Appli
28 70 100.0 585 15 US-10-433-575A-11 Sequence 2, Appli
29 70 100.0 585 15 US-10-432-262-26 Sequence 11, Appl
30 70 100.0 604 10 US-09-984-010-7 Sequence 26, Appl
31 70 100.0 609 10 US-09-919-039-370 Sequence 370, App
32 70 100.0 609 12 US-10-609-346-12 Sequence 12, Appl
33 70 100.0 609 13 US-10-153-064-7 Sequence 7, Appli
34 70 100.0 609 14 US-10-153-604A-7 Sequence 7, Appli
35 70 100.0 609 14 US-10-153-623-23 Sequence 23, Appl
36 70 100.0 610 9 US-09-984-186-2 Sequence 2, Appli
37 70 100.0 610 14 US-10-237-667-2 Sequence 2, Appli
38 70 100.0 610 14 US-10-237-708-2 Sequence 2, Appli
39 70 100.0 610 14 US-10-237-866-2 Sequence 2, Appli
40 70 100.0 610 14 US-10-237-871-2 Sequence 2, Appli
41 70 100.0 610 14 US-10-237-824-2 Sequence 2, Appli
42 70 100.0 616 12 US-10-433-108-13 Sequence 13, Appl
43 70 100.0 624 12 US-10-433-108-16 Sequence 16, Appl
44 70 100.0 631 12 US-10-433-108-14 Sequence 14, Appl
45 70 100.0 640 12 US-10-433-108-15 Sequence 15, Appl

```

ALIGNMENTS

```

RESULT 1
US-10-074-956-24
; Sequence 24, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-956-24

```

Query Match 100.0%; Score 70; DB 13; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TVATLRETYGEMAD 14
Db 76 TVATLRETYGEMAD 89

```

```

RESULT 2
US-10-074-956-27
; Sequence 27, Application US/10074956
; Publication No. US20020193332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956

```

```

XX 16-JUL-2001; 2001WO-US022263.
XX
XX 14-JUL-2000; 2000US-0218381P.
XX 18-AUG-2000; 2000US-0226382P.
XX 06-OCT-2000; 2000US-0238380P.
XX 29-DEC-2000; 2000US-0258764P.
XX 14-JUN-2001; 2001US-0298317P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX WPI; 2002-195801/25.
XX
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
XX stimulating hormone concatamer or its analog, for treating inflammatory
XX or autoimmune disorders.
XX
XX Example 2; Page 4-5; 89pp; English.
XX
XX The present invention relates to a nucleic acid comprising a sequence
XX encoding a fusion polypeptide having an alpha-melanocyte stimulating
XX hormone (MSH) concatamer. The sequences are useful for treating an
XX individual suffering from, or at risk of, a disorder of the immune system
XX e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
XX arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
XX hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
XX multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
XX present sequence is a protein described in the exemplification of the
XX invention
XX
XX Sequence 241 AA;
XX
XX Query Match 100.0%; Score 70; DB 5; Length 241;
XX Best Local Similarity 100.0%; Pred. No. 0.0001;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TVATLRETYGEMAD 14
XX |||||
XX Db 100 TVATLRETYGEMAD 113
XX
XX Search completed: April 19, 2004, 11:51:17
XX Job time : 9.23453 secs

```

DE Novel human secreted protein #365.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 OS
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 XX 25-OCT-2001.
 XX
 XX 16-APR-2001; 2001WO-US008656.
 XX
 XX 18-APR-2000; 2000US-00552929.
 XX
 XX 26-JAN-2001; 2001US-00770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 PI
 XX WPI; 2001-611725/70.
 DR
 XX Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 PT
 XX
 PS Claim 20; Page 205; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 214 AA;
 Query Match 100.0%; Score 70; DB 4; Length 214;
 Best Local Similarity 100.0%; Pred. No. 8.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TVATLRETYGEMAD 14
 |||||
 DB 112 TVATLRETYGEMAD 125
 RESULT 14
 AA017051
 ID AA017051 standard; protein; 236 AA.
 XX
 AC AA017051;
 XX
 DT 29-MAY-2002 (first entry)
 XX
 XX Human albumin-thrombin-alpha-MSH SEQ ID NO: 82.
 DE
 XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
 KW immunosuppressive; antineoplastic; antirheumatic; antiarthritic;
 KW antidiabetic; antibacterial; dermatological; antipsoriatic;
 KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
 XX diabetes; uveitis; coeliac disease.

KW diabetes; uveitis; coeliac disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200206316-A2.
 XX
 PD 24-JAN-2002.
 XX
 XX 16-JUL-2001; 2001WO-US022263.
 XX
 XX 14-JUL-2000; 2000US-0218381P.
 XX
 PR 18-AUG-2000; 2000US-0226382P.
 PR
 PR 06-OCT-2000; 2000US-0238380P.
 PR
 PR 29-DEC-2000; 2000US-0258764P.
 PR
 PR 14-JUN-2001; 2001US-0298317P.
 XX
 PA (ZYCO-) ZYCO INC.
 XX
 XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
 PI
 XX WPI; 2002-195801/25.
 DR
 XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
 PT stimulating hormone concatamer or its analog, for treating inflammatory
 PT or autoimmune disorders.
 PT
 XX
 PS Example 2; Page 48; 89pp; English.
 XX
 CC The present invention relates to a nucleic acid comprising a sequence
 CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
 CC hormone (MSH) concatamer. The sequences are useful for treating an
 CC individual suffering from, or at risk of, a disorder of the immune system
 CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
 CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
 CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
 CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
 CC present sequence is a peptide described in the exemplification of the
 CC invention
 XX
 SQ Sequence 236 AA;
 Query Match 100.0%; Score 70; DB 5; Length 236;
 Best Local Similarity 100.0%; Pred. No. 9.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TVATLRETYGEMAD 14
 |||||
 DB 100 TVATLRETYGEMAD 113
 RESULT 15
 AA016984
 ID AA016984 standard; protein; 241 AA.
 XX
 AC AA016984;
 XX
 XX 29-MAY-2002 (first entry)
 DT
 XX Alpha-MSH construct protein fragment SEQ ID NO: 59.
 DE
 XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
 KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
 KW immunosuppressive; antineoplastic; antirheumatic; antiarthritic;
 KW antidiabetic; antibacterial; dermatological; antipsoriatic;
 KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
 XX diabetes; uveitis; coeliac disease.
 XX Unidentified.
 XX
 PN WO200206316-A2.
 XX
 XX 24-JAN-2002.

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
 Db 76 TVATLRETYGEMAD 89

RESULT 11

ABU10022
 ID ABU10022 standard; protein; 195 AA.

AC ABU10022;

DT 31-JUL-2003 (first entry)

DE Human serum albumin residues 1-195.

KW Bladder disorder; cytostatic; antiinflammatory; immune response;
 KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;
 KW bladder cancer; tumour; interstitial cystitis; inflammation;
 KW alpha-MSH concatamer; melanocyte stimulating hormone; human;
 KW serum albumin.

XX Homo sapiens.

XX US2002193332-A1.

XX 19-DEC-2002.

XX 12-FEB-2002; 2002US-00074956.

XX 12-FEB-2001; 2001US-0268175P.

XX (HEDL/) HEDLEY M L.

XX Hedley ML;

XX WPI; 2003-447327/42.

PT Modulating immune responses in a mammal with a bladder disorder e.g.
 PT bladder cancer, by administering nucleic acids comprising un-methylated
 PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to
 PT the mammal.

PS Example 2; Page 9; 17pp; English.

XX The invention describes a method of modulating an immune response in a
 CC mammal, comprising identifying a mammal that has or is at risk for having
 CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)
 CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated
 CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;
 CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The
 CC method is useful for modulating immune response in a mammal having a
 CC bladder disorder, where administration of (N1) results in an amelioration
 CC of one or more symptoms of the disorder. Preferably, the method is useful
 CC for modulating immune response in a mammal having bladder cancer (where
 CC administration of (N1) results in a decrease in tumour size or activity),
 CC or for modulating immune response in a mammal having interstitial
 CC cystitis (where administration of (N1) results in a modulation of the
 CC immune response from T2 response to a T1 response). The method is also
 CC useful for modulating immune response in a mammal having bladder disorder
 CC that is characterised by inflammation which is associated with symptoms
 CC of interstitial cystitis or associated with a disruption of the integrity
 CC of the bladder lining. This is the amino acid sequence of human serum
 CC albumin residues 1-195 that can be used in the creation of melanocyte
 CC stimulating hormone (alpha-MSH) concatamers resulting in secretion of the
 CC fusion protein when expressed in mammalian cells

XX Sequence 195 AA;

Query Match 100.0%; Score 70; DB 7; Length 195;

Best Local Similarity 100.0%; Pred. No. 7.9e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
 Db 76 TVATLRETYGEMAD 89

RESULT 12

AAV83947
 ID AAV83947 standard; protein; 204 AA.

AC AAV83947;

DT 28-JUL-2000 (first entry)

DE Yeast codon-biased recombinant HSA protein fragment HSA-I.

KW Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
 KW overlapping oligonucleotide; expression vector.

XX Homo sapiens.

XX Synthetic.

XX CN1239103-A.

XX -22-DEC-1999.

XX 17-JUN-1998; 98CN-00102506.

XX 17-JUN-1998; 98CN-00102506.

XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.

XX Li S, Lu D;

XX WPI; 2000-351198/31.

XX N-PSDB; AAA10092.

PT Process for preparing recombinant human serum albumin comprising yeast
 PT biased sex codons - uses a recombinant DNA technique.

PS Example 1; Fig 3; 44pp; Chinese.

XX The method relates to a method of recombinantly producing human serum
 CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
 CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
 CC three synthetic fragments (AAA10092-A10094) joined by recombinant DNA
 CC technology. Each HSA fragment was synthesised from overlapping
 CC oligonucleotide fragments that were extended. This sequence represents
 CC the sequence of the HSA fragment HSA-I encoded by the human gene with a
 CC yeast codon bias. The invention also covers a recombinant expression
 CC vector, yeast host cells carrying the recombinant expression vector and
 CC the process for producing human serum albumin in the yeast host cell,
 CC especially in secretory mode

XX Sequence 204 AA;

Query Match 100.0%; Score 70; DB 3; Length 204;

Best Local Similarity 100.0%; Pred. No. 8.3e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TVATLRETYGEMAD 14
 Db 83 TVATLRETYGEMAD 96

RESULT 13

AAU29874
 ID AAU29874 standard; protein; 214 AA.

AC AAU29874;

DT 18-DEC-2001 (first entry)

XX

CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 156 AA;

Query Match 100.0%; Score 70; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 6.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
| | | | | | | | | | | | | | |
Db 31 TVATLRETYGEMAD 44

RESULT 9
AAU33271
ID AAU33271 standard; protein; 156 AA.

AC AAU33271;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #3762.

XX Human; vaccination; Gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US008656.

PR 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.

PS Claim 20; Page 751; 765pp; English.

XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.

CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
SQ Sequence 156 AA;

Query Match 100.0%; Score 70; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 6.1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
| | | | | | | | | | | | | | |
Db 31 TVATLRETYGEMAD 44

RESULT 10
AAO17048
ID AAO17048 standard; protein; 195 AA.

AC AAO17048;

XX 29-MAY-2002 (first entry)

XX Human serum albumin (1-195) SEQ ID NO: 56.

XX Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritis;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.

XX Homo sapiens.

XX WO200206316-A2.

XX 24-JAN-2002.

XX 16-JUL-2001; 2001WO-US022263.

XX 14-JUL-2000; 2000US-0218381P.

XX 18-AUG-2000; 2000US-0226382P.

XX 06-OCT-2000; 2000US-0238380P.

XX 29-DEC-2000; 2000US-0258764P.

XX 14-JUN-2001; 2001US-0298317P.

XX (ZYCO-) ZYCOS INC.

XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;

XX WPI; 2002-195801/25.

XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.

XX Example 2; Page 46; 89pp; English.

XX The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a protein described in the exemplification of the
CC invention
XX
SQ Sequence 195 AA;

Query Match 100.0%; Score 70; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 7.9e-05;

```

XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR N-PSDB; AA182561.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 20; SEQ ID NO 16522; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AA179941-AA193841) and
XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 124 AA;
    Query Match 100.0%; Score 70; DB 4; Length 124;
    Best Local Similarity 100.0%; Pred. No. 4.7e-05;
    Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGENAD 14
DB 96 TVATLRETYGENAD 109

RESULT 7
AAU29708
ID AAU29708 standard; protein; 134 AA.
XX AC AAU29708;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #199.
XX KW Human; vaccination; gene therapy; nutritional supplement;
XX KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX PN WO200179449-A2.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-US008656.
XX PR 18-APR-2000; 2000US-00552929.
XX PR 26-JAN-2001; 2001US-00770160.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-611725/70.
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX PT vaccination, testing and therapy.
XX PS Claim 20; Page 183; 765pp; English.

XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-611725/70.
XX PT Nucleic acids encoding a range of human polypeptides, useful for
XX PT determining the presence of or predisposition to a disease associated
XX CC with altered levels of polypeptide. The polypeptides are also useful for
XX CC identifying agents (agonists and antagonists) that bind to them. Cells
XX CC expressing the proteins are useful for identifying a therapeutic agent
XX CC for use in treatment of a pathology related to aberrant expression or
XX CC physiological interactions of the polypeptide. Vectors comprising the
XX CC nucleic acids encoding the polypeptides and cells genetically engineered
XX CC to express them are also useful for producing the proteins. The proteins
XX CC are useful in genetic vaccination, testing and therapy, and can be used
XX CC as nutritional supplements. They may be used to increase stem cell
XX CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
XX CC and/or nerve tissue growth or regeneration; immune suppression and/or
XX CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
XX CC AAU29510-AAU33304 represent the amino acid sequences of novel human
XX CC secreted proteins of the invention
XX SQ Sequence 134 AA;
    Query Match 100.0%; Score 70; DB 4; Length 134;
    Best Local Similarity 100.0%; Pred. No. 5.2e-05;
    Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGENAD 14
DB 7 TVATLRETYGENAD 20

RESULT 8
AAU33073
ID AAU33073 standard; protein; 156 AA.
XX AC AAU33073;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human secreted protein #3564.
XX KW Human; vaccination; gene therapy; nutritional supplement;
XX KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX OS Homo sapiens.
XX PN WO200179449-A2.
XX PD 25-OCT-2001.
XX PF 16-APR-2001; 2001WO-US008656.
XX PR 18-APR-2000; 2000US-00552929.
XX PR 26-JAN-2001; 2001US-00770160.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-611725/70.
XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX PT vaccination, testing and therapy.
XX PS Claim 20; Page 704; 765pp; English.
XX CC The invention relates to novel human secreted polypeptides. The
XX CC polypeptides and antibodies to the polypeptides are useful for
XX CC determining the presence of or predisposition to a disease associated
XX CC with altered levels of polypeptide. The polypeptides are also useful for
XX CC identifying agents (agonists and antagonists) that bind to them. Cells
XX CC expressing the proteins are useful for identifying a therapeutic agent
XX CC for use in treatment of a pathology related to aberrant expression or

```

```

XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 18316.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.
XX FN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR N-PSDB; AAI84355.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 20; SEQ ID NO 18316; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 123 AA;
XX Query Match 100.0%; Score 70; DB 4; Length 123;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 98 TVATLRETYGEMAD 111

RESULT 5
AAO04423
ID AAO04423 standard; protein; 123 AA.
XX AC AAO04423;
XX OS Homo sapiens.
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 18315.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX
```

```

OS Homo sapiens.
XX FN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR N-PSDB; AAI84354.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 20; SEQ ID NO 18315; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 123 AA;
XX Query Match 100.0%; Score 70; DB 4; Length 123;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-05;
XX Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 99 TVATLRETYGEMAD 112

RESULT 6
AAO02630
ID AAO02630 standard; protein; 124 AA.
XX AC AAO02630;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 16522.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.
XX FN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX
```

CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 114 AA;

Query Match 100.0%; Score 70; DB 4; Length 114;
Best Local Similarity 100.0%; Pred. No. 4.3e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 101 TVATLRETYGEMAD 114

RESULT 2
AAO04446
ID AAO04446 standard; protein; 118 AA.

XX AC AAO04446;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 18338.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AAI84377.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing

XX PT and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 20; SEQ ID NO 18338; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 118 AA;

Query Match 100.0%; Score 70; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 98 TVATLRETYGEMAD 111

RESULT 3

AAO01996

ID AAO01996 standard; protein; 119 AA.

XX AC AAO01996;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 15888.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US004927.

XX PR 28-FEB-2000; 2000US-00515126.

XX PR 18-MAY-2000; 2000US-00577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-514838/56.

XX DR N-PSDB; AAI81927.

XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing

XX PT and treating e.g. leukemia, inflammation and immune disorders.

XX PS Claim 20; SEQ ID NO 15888; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 119 AA;

Query Match 100.0%; Score 70; DB 4; Length 119;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
DB 98 TVATLRETYGEMAD 111

RESULT 4

AAO04424

ID AAO04424 standard; protein; 123 AA.

XX AC AAO04424;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 8.23453 Seconds
(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_76_89

Perfect score: 70

Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseqpl980s.*
- 2: Geneseqpl990s.*
- 3: Geneseqpl2000s.*
- 4: Geneseqpl2001s.*
- 5: Geneseqpl2002s.*
- 6: Geneseqpl2003as.*
- 7: Geneseqpl2003bs.*
- 8: Geneseqpl2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	70	100.0	114	AAO09588	AAO09588 Human pol
2	70	100.0	118	AAO04446	AAO04446 Human pol
3	70	100.0	119	AAO01996	AAO01996 Human pol
4	70	100.0	123	AAO04424	AAO04424 Human pol
5	70	100.0	123	AAO04423	AAO04423 Human pol
6	70	100.0	124	AAO02630	AAO02630 Human pol
7	70	100.0	134	AAU29708	AAU29708 Novel hum
8	70	100.0	156	AAU33073	AAU33073 Novel hum
9	70	100.0	156	AAU33271	AAU33271 Novel hum
10	70	100.0	195	AAO17048	AAO17048 Human ser
11	70	100.0	195	ABU10022	ABU10022 Human ser
12	70	100.0	204	AAU83947	AAU83947 Yeast cod
13	70	100.0	214	AAU29874	AAU29874 Novel hum
14	70	100.0	236	AAO17051	AAO17051 Human alb
15	70	100.0	241	AAO16984	AAO16984 Alpha-MSH
16	70	100.0	241	ABU10025	ABU10025 Alpha-MSH
17	70	100.0	242	AAO16985	AAO16985 Alpha-MSH
18	70	100.0	244	AAO16986	AAO16986 Alpha-MSH
19	70	100.0	245	AAO16987	AAO16987 Alpha-MSH
20	70	100.0	245	AAO16988	AAO16988 Alpha-MSH
21	70	100.0	268	AAO16989	AAO16989 Alpha-MSH
22	70	100.0	268	ABU10026	ABU10026 Alpha-MSH
23	70	100.0	303	AAU14178	AAU14178 Human ser
24	70	100.0	327	AAU32564	AAU32564 Novel hum
25	70	100.0	327	AAU29942	AAU29942 Novel hum

ALIGNMENTS

RESULT 1

AAO09588
ID AAO09588 standard; protein; 114 AA.
XX
AC AAO09588;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 23480.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
XX
PA 18-MAY-2000; 2000US-00577409.
(HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-514838/56.
DR N-PSDB; AAI89519.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 20; SEQ ID NO 23480; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO1910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

26 70 100.0 327 4 AAU32995
27 70 100.0 327 4 AAU3287
28 70 100.0 373 1 AAP90387
29 70 100.0 388 1 AAP90389
30 70 100.0 389 1 AAP90390
31 70 100.0 390 1 AAP90391
32 70 100.0 401 4 AAU29876
33 70 100.0 407 1 AAP90392
34 70 100.0 500 7 ADD32019
35 70 100.0 584 6 ABG72381
36 70 100.0 585 1 AAP93344
37 70 100.0 585 1 AAP90388
38 70 100.0 585 1 AAP91422
39 70 100.0 585 2 AAR05318
40 70 100.0 585 2 AAR08457
41 70 100.0 585 2 AAR26207
42 70 100.0 585 2 AAR26362
43 70 100.0 585 2 AAR20029
44 70 100.0 585 2 AAR80301
45 70 100.0 585 2 AAO20111

AAU32995 Novel hum
AAU3287 Novel hum
AAP90387 N-termina
AAP90389 N-termina
AAP90390 N-termina
AAP90391 N-termina
AAU29876 Novel hum
AAP90392 N-termina
ADD32019 Heterolog
ABG72381 Mature hu
AAP93344 Sequence
AAP90388 Mature hu
AAP91422 Human nor
AAR05318 Human ser
AAR08457 Human ser
AAR26207 Human ser
AAR26362 Synthetic
AAR20029 Human ser
AAR80301 Human ser
AAO20111 HSA prote

Search completed: April 19, 2004, 12:00:01
Job time : 7.40351 secs

RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RL Bordetella parapertussis and Bordetella bronchiseptica."
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640432; CAE38219.1; -.
 KW Complete proteome.
 SQ SEQUENCE 281 AA; 30591 MW; C97F4A51E40E3315 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 281;
 Best Local Similarity 66.7%; Pred. No. 32;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 ATLRETYGEMAD 14
 ||||| |||||
 Db 266 ATLRLYGLID 277

RESULT 13
 ID O01889 PRELIMINARY; PRT; 354 AA.
 AC O01889;
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE R08F11.4 protein.
 GN R08F11.4
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6339;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Minx P.;
 RT "The sequence of C. elegans cosmid R08F11.";
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003385; AAB54246.1; -.
 DR PIR; H89009; H89009
 DR WormPep; R08F11.4; CE12586.
 DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
 DR InterPro; IPR000051; SAM_bind.
 SQ SEQUENCE 354 AA; 39365 MW; CB422510DFA3E0A CRC64;

Query Match 58.6%; Score 41; DB 5; Length 354;
 Best Local Similarity 75.0%; Pred. No. 63;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VATLRETYGEMA 13
 ||||| |||||
 Db 280 VFTDRTYKGMA 291

RESULT 14
 ID Q9A8L1 PRELIMINARY; PRT; 692 AA.
 AC Q9A8L1;
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Glycyl-tRNA synthetase, beta subunit.
 GN CC1342.

OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Ghann M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Uitterback L., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RL "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; A3005809; AAK23323.1; -.
 DR PIR; G87415; G87415.
 DR TIGR; CC1342; -.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004820; F:glycine-tRNA ligase activity; IEA.
 DR GO; GO:0006426; P:glycyl-tRNA aminoacylation; IEA.
 DR InterPro; IPR002311; tRNA synt 2f.
 DR InterPro; IPR006194; tRNA synt Gly.
 DR Pfam; PF02092; tRNA synt 2f; 1.
 DR PRINTS; PR01045; TENASYNTHGB.
 DR TIGRFAMs; TIGR00211; glys; 1.
 DR PROSITE; PFS0861; AA-TRNA-LIGASE-II_GLYAB; 1.
 KW Aminoacyl-tRNA synthetase; Complete proteome.
 SQ SEQUENCE 692 AA; 76060 MW; 3CDE4E730D097846 CRC64;

Query Match 58.6%; Score 41; DB 16; Length 692;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
 ||||| |||||
 Db 673 TLATVRDAMGVAD 686

RESULT 15
 ID Q82G98 PRELIMINARY; PRT; 169 AA.
 AC Q82G98;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN SAV4000.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.";


```

DB 01-MAR-2002 (TReMBLrel. 20, Created)
DB 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DB 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DB Hypothetical protein A115024.
DB ALU5024.
DB Anabaena sp. (strain PCC 7120).
DB Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
DB NCBI_TaxID=103690;
DB [1]
DB SEQUENCE FROM N.A.
DB MEDLINE=21595285; PubMed=11759840;
DB Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
DB Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
DB Kishida Y., Kohara M., Matsunoto M., Matsuno A., Muraki A.,
DB Nakazaki N., Shampo S., Sugimoto M., Takazawa M., Yamada M.,
DB Yasuda M., Tabata S.;
DB "Complete genomic sequence of the filamentous nitrogen-fixing
DB cyanobacterium Anabaena sp. strain PCC 7120.";
DB DNA Res. 8:205-213(2001);
DB EMBL; AP003598; BAB76723.1; -.
DB PIR; AH2433;
DB Hypothetical protein; Complete proteome.
DB SEQUENCE 252 AA; 29440 MW; 3ABDCE406345B50F CRC64;

Query Match 60.0%; Score 42; DB 16; Length 252;
Best Local Similarity 64.3%; Pred. No. 29;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TVATLRITYGEMAD 14
:|||||
Db 176 TIAIIRSTAHMMD 189

RESULT 10
QWIF5 PRELIMINARY; PRT; 271 AA.
AC Q7WIF5;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE Putative exported protein.
GN B2896.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR EMBL; EX640414; CAE41588.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29614 MW; E75E8C1C2337AB06 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||
Db 256 ATLRLYGELID 267

RESULT 12
QW613 PRELIMINARY; PRT; 281 AA.
AC Q7W613;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE Putative exported protein.
GN BPP2926.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR EMBL; EX640445; CAE33388.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29643 MW; FC5E8C19A837AB0E CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||

```

```

DB 256 ATLRLYGELID 267

RESULT 11
Q7VM6 PRELIMINARY; PRT; 271 AA.
AC Q7VM6;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative exported protein.
GN BPI292.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR EMBL; EX640414; CAE41588.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29614 MW; E75E8C1C2337AB06 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||
Db 256 ATLRLYGELID 267

RESULT 12
QW613 PRELIMINARY; PRT; 281 AA.
AC Q7W613;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE Putative exported protein.
GN BPP2926.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR EMBL; EX640414; CAE41588.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29614 MW; E75E8C1C2337AB06 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||
Db 256 ATLRLYGELID 267

RESULT 12
QW613 PRELIMINARY; PRT; 281 AA.
AC Q7W613;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DE Putative exported protein.
GN BPP2926.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
DR EMBL; EX640414; CAE41588.1; -.
KW Complete proteome.
SQ SEQUENCE 271 AA; 29614 MW; E75E8C1C2337AB06 CRC64;

Query Match 60.0%; Score 42; DB 16; Length 271;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ATLRETYGEMAD 14
|||||

```

Best Local Similarity 61.5%; Pred. No. 13; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 3

QY 2 VATLRETYGEMAD 14
DB 69 IPNLRENYGELAD 81

RESULT 6

Q8C7H3 PRELIMINARY; PRT; 608 AA.
AC Q8C7H3
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Albumin 1.
DE ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005396; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum_albumin; 1.
DR SMART; SMC0103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR SEQUENCE 608 AA; 68722 MW; 292P600BED3A61B4 CRC64;

Query Match 65.7%; Score 46; DB 11; Length 608;

Best Local Similarity 61.5%; Pred. No. 14; Indels 0; Gaps 0;
Matches 8; Conservative 2; Mismatches 3

QY 2 VATLRETYGEMAD 14
DB 101 IPNLRENYGELAD 113

RESULT 7

Q8C9N0 PRELIMINARY; PRT; 116 AA.
AC Q8C9N0
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Pan hematopoietic expression.
DE PHEMX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).
DR EMBL; AK041748; BAC31052.1; -.
DR MGD; MGI:1350360; PHEMX.
DR GO; GO:0005622; C:intracellular; IEA.
SQ SEQUENCE 116 AA; 12587 MW; 7638C7C4CFD2CDDC CRC64;

Query Match 60.0%; Score 42; DB 11; Length 116;
Best Local Similarity 61.5%; Pred. No. 13; Indels 0; Gaps 0;
Matches 8; Conservative 3; Mismatches 2

QY 1 TVATLRETYGEMA 13
DB 80 TIATVREAGHGLMA 92

RESULT 8

Q7TNQ7 PRELIMINARY; PRT; 201 AA.
AC Q7TNQ7
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haeh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Krzywinski M.I., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; SC055858; AAH55858.1; -.
KW Hypothetical protein.
SQ SEQUENCE 201 AA; 22722 MW; 09421AC54204F79B CRC64;

Query Match 60.0%; Score 42; DB 11; Length 201;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMA 13
DB 53 TIATVREAGHGLMA 65

RESULT 9

Q8YMB0 PRELIMINARY; PRT; 252 AA.
ID Q8YMB0
AC Q8YMB0;

```

ID QYSG3 PRELIMINARY; PRT; 584 AA.
AC QYSG3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin (Fragment).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
RA Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
RT lymphoproliferative responses in atopic patients.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ487677; CAD32275.1; -.
DR NON_TER 1
FT SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 81.4%; Score 57; DB 6; Length 584;
Best Local Similarity 78.6%; Pred. No. 0.15;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVATLRITYGEMAD 14
Db 76 TVASLRDXYGEMAD 89

RESULT 3
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Stadecker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RT component in murine infection with Schistosoma mansoni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF418550; AAL08579.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EAB28E1C66E54 CRC64;

Query Match 70.0%; Score 49; DB 5; Length 608;
Best Local Similarity 61.5%; Pred. No. 4.2;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 VATLRITYGEMAD 14
Db 101 IPTLRDSYGLAD 113

RESULT 4
```

```

Q8CG74 PRELIMINARY; PRT; 205 AA.
ID Q8CG74;
AC Q8CG74;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvetOla;
RA Van Reeth T., Dreze P.L., Gabant P., Szpirer C., Szpirer J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ277794; CAC81903.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 1.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 1.
DR NON_TER 205
FT SEQUENCE 205 AA; 23624 MW; 0A982E85C7474BF9 CRC64;
SQ SEQUENCE 205 AA; 23624 MW; 0A982E85C7474BF9 CRC64;

Query Match 65.7%; Score 46; DB 11; Length 205;
Best Local Similarity 61.5%; Pred. No. 4.5;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VATLRITYGEMAD 14
Db 101 IPTLRDSYGLAD 113

RESULT 5
Q8C7C7 PRELIMINARY; PRT; 576 AA.
ID Q8C7C7;
AC Q8C7C7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL: AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Albl.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR NON_TER 1
FT SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;
SQ SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;

Query Match 65.7%; Score 46; DB 11; Length 576;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 5.40351 Seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_76_89

Perfect score: 70

Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	396	4 Q8IU7	Q8IU7 homo sapien
2	57	81.4	584	6 Q7YSG3	Q7YSG3 felis silve
3	49	70.0	608	5 Q95VB7	Q95VB7 schistosoma
4	46	65.7	205	11 Q8CG74	Q8CG74 mus musculus
5	46	65.7	576	11 Q8C7C7	Q8C7C7 mus musculus
6	46	65.7	608	11 Q8C7H3	Q8C7H3 mus musculus
7	42	60.0	116	11 Q8C9N0	Q8C9N0 mus musculus
8	42	60.0	201	11 Q7TNQ7	Q7TNQ7 mus musculus
9	42	60.0	252	16 Q8YMB0	Q8YMB0 anabaena sp
10	42	60.0	271	16 Q7WIF5	Q7WIF5 bordetella
11	42	60.0	271	16 Q7VYM6	Q7VYM6 bordetella
12	42	60.0	281	16 Q7W6I3	Q7W6I3 bordetella
13	41	58.6	354	5 O01889	O01889 caenorhabd
14	41	58.6	692	16 Q9A8L1	Q9A8L1 caulobacter
15	40	57.1	169	16 Q82G98	Q82G98 streptomyce
16	40	57.1	270	16 Q8F5F7	Q8F5F7 leptospira

```

17 40 57.1 314 16 Q83BP5
18 40 57.1 330 16 Q97R21
19 40 57.1 330 16 Q8E6D5
20 40 57.1 330 16 Q8E6D5
21 40 57.1 330 16 Q8E6D5
22 40 57.1 563 5 Q964I3
23 40 57.1 585 16 Q9RR13
24 40 57.1 924 2 Q92IG3
25 40 57.1 924 2 Q9X582
26 39.5 56.4 202 2 Q9ZGF5
27 39 55.7 157 5 Q90743
28 39 55.7 176 16 Q8DL82
29 39 55.7 179 16 Q9CELO
30 39 55.7 250 5 Q8SRK2
31 39 55.7 443 2 Q913B5
32 39 55.7 443 2 Q53162
33 39 55.7 444 2 Q9X7L9
34 39 55.7 445 16 Q8P181
35 39 55.7 541 16 Q9RKD1
36 39 55.7 633 12 Q92095
37 39 55.7 1049 2 Q9XBP6
38 39 55.7 3898 12 Q92365
39 39 55.7 3898 12 Q92872
40 38 54.3 91 15 Q8XSP0
41 38 54.3 124 17 Q8ZVT1
42 38 54.3 251 16 Q8XFB9
43 38 54.3 308 16 Q8XRG7
44 38 54.3 318 2 Q7WZ93
45 38 54.3 378 16 Q8FPK8

```

ALIGNMENTS

RESULT 1

```

Q8IU7
ID Q8IU7 PRELIMINARY; PRT; 396 AA.
AC Q8IU7;
DT 01-VAR-2003 (TRENBLrel. 23, Created)
DT 01-VAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1; -
DR GO; GO:0005655; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

```

Query Match 100.0%; Score 70; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14

DB 100 TVATLRETYGEMAD 113

RESULT 2

Q7YSG3

Query Match 57.1%; Score 40; DB 1; Length 409;
 Best Local Similarity 70.0%; Pred. No. 7.8;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 VAIURETYGE 11
 : |||: |||
 Db 217 INTLRKTYGE 226

Search completed: April 19, 2004, 11:52:48
 Job time : 3.02124 secs

DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glycine dehydrogenase [decarboxylating], mitochondrial precursor
 DE (EC 1.4.4.2) (Glycine decarboxylase) (Glycine cleavage system P-
 protein)
 DE GCV2 OR GSD2 OR YMR189W OR YM9646.01.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_taxid=94932;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96120340; PubMed=7498764;
 RA Sinclair D.A., Dawes I.W.;
 RT "Genetics of the synthesis of serine from glycine and the utilization
 of glycine as sole nitrogen source by *Saccharomyces cerevisiae*.";
 RL Genetics 140:1213-1222(1995).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=S288c / AB972;
 RX MEDLINE=97313268; PubMed=9169872;
 RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
 RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
 RA Jagsels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
 RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome
 XIII.";
 RL Nature 387:90-93(1997).
 CC -!- FUNCTION: The glycine cleavage system catalyzes the degradation of
 glycine. The P-protein binds the alpha-amino group of glycine
 through its pyridoxal phosphate cofactor; CO(2) is released and
 the remaining methylamine moiety is then transferred to the
 lipamide cofactor of the H-protein.
 CC -!- CATALYTIC ACTIVITY: Glycine + lipoylprotein = S-
 aminomethylidihydrolipoylprotein + CO(2).
 CC -!- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC -!- SIMILARITY: Belongs to the gcvp family.

RESULT 15

OD02_HAEIN
 ID ODO2_HAEIN STANDARD; PRT; 409 AA.
 AC P45302;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Dihydrolipoamide succinyltransferase component of 2-oxoglutarate
 DE dehydrogenase complex (EC 2.3.1.61) (E2).
 GN SUCB OR H11661.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kierlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.W.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 Rd.";
 RL Science 269:496-512(1995).
 CC -!- FUNCTION: THE 2-OXOGLUTARATE DEHYDROGENASE COMPLEX CATALYZES THE
 OVERALL CONVERSION OF 2-OXOGLUTARATE TO SUCCINYL-COA & CO(2). IT
 CONTAINS MULTIPLE COPIES OF 3 ENZYMIC COMPONENTS: 2-OXOGLUTARATE
 DEHYDROGENASE (E1), DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE (E2) AND
 LIPOAMIDE DEHYDROGENASE (E3) (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: Succinyl-CoA + dihydrolipoamide = CoA + S-
 succinylidihydrolipoamide.
 CC -!- COFACTOR: Contains 1 covalently bound lipoyl cofactor (By
 similarity).
 CC -!- PATHWAY: Tricarboxylic acid cycle.
 CC -!- SUBUNIT: Forms a 24-polypeptide structural core with octahedral
 symmetry (By similarity).
 CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
 CC -!- SIMILARITY: Contains 1 lipoyl-binding domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U32839; AAC23307.1; -.
 DR PIR; D64135; D64135.
 DR KSSP; P07016; 1C4T.
 DR TIGR; H11661; -.
 DR InterPro; IPR001078; 2Oxoacid dh.
 DR InterPro; IPR000089; Biotin lipoyl.
 DR InterPro; IPR004167; E3_binding.
 DR InterPro; IPR003016; Lipoy_BS.
 DR InterPro; IPR006255; SUCB.
 DR Pfam; PF00198; 2-oxoacid dh; 1.
 DR Pfam; PF02817; biotin lipoyl; 1.
 DR Pfam; PF00364; e3 binding; 1.
 DR ProDom; PD001115; 2Oxoacid dh; 1.
 DR TIGRFAMs; TIGR01347; sucB; 1.
 DR PROSITE; PS00189; LIPOYL; 1.
 DR Tricarboxylic acid cycle; transferase; Acyltransferase; Lipoyl;
 KW Complete proteome.
 FT BINDING 43 43 LIPOYL (POTENTIAL).
 FT ACT_SITE 380 380 POTENTIAL.
 FT ACT_SITE 384 384 POTENTIAL.
 SQ SEQUENCE 409 AA; 45162 MW; 3FBE62BC17433839 CRC64;

Query Match 58.6%; Score 41; DB 1; Length 1034;
 Best Local Similarity 66.7%; Pred. No. 14;
 Matches 8; Conservative 2; Mismatches 0; Gaps 0;
 QY 1 TVATLREYIGEM 12
 ||| :|||:
 DB 1006 TVARLDYIGDM 1017

```

FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
FT VARIANT 262 262
FT CONFLICT 174 174
SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;

Query Match 60.0%; Score 42; DB 1; Length 608;
Best Local Similarity 53.8%; Pred. No. 5.1;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VATIRETYGEMAD 14
: |||: |||: |||
Db 101 IPLKRDNYGELAD 113

RESULT 13
ALBU CHICK STANDARD; PRT; 615 AA.
AC P1312;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).
GN ALB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Casady A.I.; Salkild C.K.; Baverstock P.; Wallace J.C.;
RN Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=83161037; PubMed=6187737;
RA Hache R.J.G.; Wiskocil R.; Vasa M.; Roy R.N.; Lau P.C.K.; Deeley R.G.;
RT "The 5' noncoding and flanking regions of the avian very low density
RT apolipoprotein II and serum albumin genes. Homologies with the egg
RT white protein genes.";
RL J. Biol. Chem. 258:4556-4564 (1983).
RN [3]
RN SEQUENCE OF 19-30.
RX MEDLINE=78019943; PubMed=911327;
RA Rosen A.M.; Geller D.M.;
RT "Chicken microsomal albumin: amino terminal sequence of chicken
RT proalbumin.";
RL Blochem. Biophys. Res. Commun. 78:1060-1066 (1977).
RN [4]
RN ALLERGENIC PROPERTIES.
RX MEDLINE=21381307; PubMed=11488669;
RA Quirce S.; Maranon F.; Umpierrez A.; de las Heras M.;
RA Fernandez-Caldas E.; Sastre J.;
RT "Chicken serum albumin (Gal d 5*) is a partially heat-labile inhalant
RT and food allergen implicated in the bird-egg syndrome.";
RL Allergy 56:754-762 (2001).

```

```

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IGE.
CC partially heat-labile allergen that may cause both respiratory and
CC food-allergy symptoms in patients with the bird-egg syndrome.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60688; CAA43098.1; -.
CC EMBL; V00381; CAA23680.1; -.
CC PIR; S15571; ABCHS.
CC HSP; P02768; 1E7B.
CC InterPro: IPR000264; Serum albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 23
FT CHAIN 24 615 SERUM ALBUMIN.
FT DOMAIN 24 209 ALBUMIN 1.
FT DOMAIN 216 401 ALBUMIN 2.
FT DOMAIN 408 599 ALBUMIN 3.
FT METAL 30 30 COPPER (BY SIMILARITY).
FT DISULFID 80 89 BY SIMILARITY.
FT DISULFID 102 118 BY SIMILARITY.
FT DISULFID 117 128 BY SIMILARITY.
FT DISULFID 152 197 BY SIMILARITY.
FT DISULFID 196 205 BY SIMILARITY.
FT DISULFID 228 274 BY SIMILARITY.
FT DISULFID 273 281 BY SIMILARITY.
FT DISULFID 293 307 BY SIMILARITY.
FT DISULFID 306 317 BY SIMILARITY.
FT DISULFID 344 389 BY SIMILARITY.
FT DISULFID 388 397 BY SIMILARITY.
FT DISULFID 420 466 BY SIMILARITY.
FT DISULFID 465 476 BY SIMILARITY.
FT DISULFID 489 505 BY SIMILARITY.
FT DISULFID 504 515 BY SIMILARITY.
FT DISULFID 542 587 BY SIMILARITY.
FT DISULFID 586 595 BY SIMILARITY.
FT CARBOHYD 500 500 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 24 24 F -> M (IN REF. 3).
SQ SEQUENCE 615 AA; 69918 MW; E59E4BBCAEC066C6 CRC64;

Query Match 58.6%; Score 41; DB 1; Length 615;
Best Local Similarity 61.5%; Pred. No. 7.9;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VATIRETYGEMAD 14
: |||: |||: |||
Db 104 VEKLADSYGAMAD 116

RESULT 14
GCSP_YEAST
ID GCSP_YEAST STANDARD; PRT; 1034 AA.
AC P49095;
DT 01-FEB-1996 (Rel. 33, Created)

```

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X12422; CAA30970.1; -;
EMBL; M36787; AAA30988.1; -;
PIR; S01382; ABPGS.
HSP; P02768; 187H.
InterPro; IPR000264; Serum_albumin.
Pfam; PF00273; transport_prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum_albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper.
NON_TER 1 1
SIGNAL <1 16 BY SIMILARITY.
PROPEP 17 22 BY SIMILARITY.
CHAIN 23 605 SERUM ALBUMIN.
DOMAIN 23 202 ALBUMIN 1.
DOMAIN 209 394 ALBUMIN 2.
DOMAIN 401 592 ALBUMIN 3.
METAL 31 31 COPPER (BY SIMILARITY).
DISULFID 75 84 BY SIMILARITY.
DISULFID 97 113 BY SIMILARITY.
DISULFID 112 123 BY SIMILARITY.
DISULFID 145 190 BY SIMILARITY.
DISULFID 189 198 BY SIMILARITY.
DISULFID 221 267 BY SIMILARITY.
DISULFID 266 274 BY SIMILARITY.
DISULFID 286 300 BY SIMILARITY.
DISULFID 299 310 BY SIMILARITY.
DISULFID 337 382 BY SIMILARITY.
DISULFID 381 390 BY SIMILARITY.
DISULFID 413 459 BY SIMILARITY.
DISULFID 458 469 BY SIMILARITY.
DISULFID 482 498 BY SIMILARITY.
DISULFID 497 508 BY SIMILARITY.
DISULFID 535 580 BY SIMILARITY.
DISULFID 579 588 BY SIMILARITY.
CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1F4FF CRC64;

Query Match 60.0%; Score 42; DB 1; Length 605;
Best Local Similarity 53.8%; Pred. No. 5;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 VATLRETYGENAD 14
: ||| |||:
Db 99 IPSLREHYGLAD 111

RESULT 12
ALBU RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN ALB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81223722; PubMed=7017712;
RA Sargent T.D., Yang M., Bonner J.;
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";

Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
[2]
RN SEQUENCE OF 1-38, AND PROCESSING.
RX MEDLINE=77249657; PubMed=893447;
RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
piece. Analysis of the direct translation product of albumin
messenger RNA.";
RT J. Biol. Chem. 252:6846-6855(1977).
[3]
RN SEQUENCE OF 25-222.
RX MEDLINE=78109429; PubMed=564345;
RA Isemura S., Ikenaka T.;
RT "Amino acid sequences of fragments I and II obtained by cyanogen
bromide cleavage of rat serum albumin.";
RL J. Biochem. 83:35-48(1978).
[4]
RN SEQUENCE OF 223-288 AND 572-608.
RX MEDLINE=76260153; PubMed=956149;
RA Isemura S., Ikenaka T.;
RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
the amino acid sequences of four fragments.";
RL J. Biochem. 79:1183-1196(1976).
[5]
RN SEQUENCE OF 166-174.
RX TISSUE-Plasma;
RC MEDLINE=87194805; PubMed=2437111;
RA Carraway R.E., Mitra S.P., Cochran D.E.;
RT "Structure of a biologically active neurotensin-related peptide
obtained from pepsin-treated albumin(s).";
RL J. Biol. Chem. 262:5968-5973(1987).
[6]
RN COPPER-BINDING.
RX MEDLINE=79001817; PubMed=80265;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Copper(II)-binding ability of human alpha-fetoprotein.";
Cancer Res. 38:3483-3486(1978).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and
blood flow (Potential).
CC -!- SUBCELLULAR LOCATION: Plasma.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; V01222; CAA24532.1; -;
PIR; A93872; ABRTS.
HSP; P02768; 1E7B.
InterPro; IPR000264; Serum_albumin.
Pfam; PF00273; transport_prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum_albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper.
SIGNAL 1 18
PROPEP 19 24 SERUM ALBUMIN.
CHAIN 25 608 NEUROTENSIN-RELATED PEPTIDE.
PEPTIDE 166 174 ALBUMIN 1.
DOMAIN 25 205 ALBUMIN 2.
DOMAIN 212 397 ALBUMIN 3.
DOMAIN 404 595

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AJ011413; CAA09617.1; -
 DR EMBL; AK010025; BAB26650.1; -
 DR EMBL; BC049971; AAA49971.1; -
 DR EMBL; M16111; AAA37190.1; -
 DR EMBL; X13060; CAA31458.1; -
 DR F01; A05139; A05139.

DR HSSP; P02768; 1E7B.

DR SWISS-2DPAGE; P07724; MOUSE.

DR MGD; MGI:87991; Alb1.

DR InterPro; IPR000264; Serum albumin.

DR Pfam; PF00273; transport prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.

FT SIGNAL 1 18

FT PROPEP 19 24

FT CHAIN 25 608

FT DOMAIN 25 205

FT DOMAIN 212 397

FT DOMAIN 404 595

FT METAL 27 27

FT METAL 27 27

FT DISULFID 77 86

FT DISULFID 99 115

FT DISULFID 114 125

FT DISULFID 148 193

FT DISULFID 192 201

FT DISULFID 224 270

FT DISULFID 269 277

FT DISULFID 289 303

FT DISULFID 302 313

FT DISULFID 340 385

FT DISULFID 384 393

FT DISULFID 416 462

FT DISULFID 461 472

FT DISULFID 485 501

FT DISULFID 500 511

FT DISULFID 538 583

FT DISULFID 592 591

FT DISULFID 27 27

FT CONFLICT 33 33

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

FT CONFLICT 41 41

RC STRAIN:New Zealand white; TISSUE:liver;

RA Sheffield W.P., Syed S., Schuyler P.D.;

RL Submitted (DEC-1994) to the EMBL/GenBank/DBAJ databases.

CC !- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC !- SUBCELLULAR LOCATION: Secreted.

CC !- TISSUE SPECIFICITY: Plasma.

CC !- SIMILARITY: Belongs to the ALB/APP/VDB family.

CC !- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U18344; AAB58347.1; -

DR HSSP; P02768; 1E7B.

DR InterPro; IPR000264; Serum albumin.

DR Pfam; PF00273; transport prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.

FT SIGNAL 1 18

FT PROPEP 19 24

FT CHAIN 25 608

FT DOMAIN 25 205

FT DOMAIN 212 397

FT DOMAIN 404 595

FT METAL 27 27

FT METAL 27 27

FT DISULFID 77 86

FT DISULFID 99 115

FT DISULFID 114 125

FT DISULFID 148 193

FT DISULFID 192 201

FT DISULFID 224 270

FT DISULFID 269 277

FT DISULFID 289 303

FT DISULFID 302 313

FT DISULFID 340 385

FT DISULFID 384 393

FT DISULFID 416 462

FT DISULFID 461 472

FT DISULFID 485 501

FT DISULFID 500 511

FT DISULFID 538 583

FT DISULFID 582 591

SQ SEQUENCE 608 AA; 68914 MW; CF5E92647AAFE9A2 CRC64;

Query Match 61.4%; Score 43; DB 1; Length 608;

Best Local Similarity 63.6%; Pred. No. 3.3;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 TIRETYGEMAD 14

Db 103 SLRDTYGDVAD 113

RESULT 10

PHMX MOUSE

ID PHMX MOUSE STANDARD; PRT: 256 AA.

AC Q5JHR2; Q923G9; Q9JHQ9; Q9JHR1; Q9JHR2; Q5JHS8;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Phenx protein (AMLI-regulated transmembrane protein 1).

GN PHMX OR TSSC6 OR ART1.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT DOMAIN 25 205 SERUM ALBUMIN.
 FT DOMAIN 212 397 ALBUMIN 1.
 FT DOMAIN 404 595 ALBUMIN 2.
 FT METAL 27 27 ALBUMIN 3.
 FT DISULFID 77 86 COPPER (BY SIMILARITY).
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 148 193 BY SIMILARITY.
 FT DISULFID 192 201 BY SIMILARITY.
 FT DISULFID 224 270 BY SIMILARITY.
 FT DISULFID 269 277 BY SIMILARITY.
 FT DISULFID 289 303 BY SIMILARITY.
 FT DISULFID 302 313 BY SIMILARITY.
 FT DISULFID 340 385 BY SIMILARITY.
 FT DISULFID 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.
 FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT CONFLICT 1 26 MKWTFISLFLEFSAYSRGLVREA -> MDT (IN REP. 2).
 FT CONFLICT 146 146 A -> R (IN REP. 2).
 FT CONFLICT 206 206 I -> T (IN REP. 2).
 FT CONFLICT 349 349 V -> A (IN REP. 2).
 FT CONFLICT 359 359 S -> A (IN REP. 1).
 FT CONFLICT 448 448 V -> V (IN REP. 5).
 FT CONFLICT 474 474 D -> E (IN REP. 1).
 SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8FF7DD8FC06 CRC64;
 Query Match 77.1%; Score 54; DB 1; Length 608;
 Best Local Similarity 71.4%; Pred. No. 0.028;
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 TVATLRETYGEMAD 14
 Db 100 TVASLRDKYGDMA 113
 RESULT 8
 ID _ALBU_MOUSE STANDARD; PRT; 608 AA.
 AC P07724; Q61802;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Serum albumin precursor.
 GN ALB OR ALB1 OR ALB-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Azawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Bairdelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RN Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Liver;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 99-516 FROM N.A.
 RX MEDLINE=88216123; PubMed=2452956;
 RA Minghetti P.P., Law S.W., Dugaiczky A.;
 RT "The rate of molecular evolution of alpha-fetoprotein approaches that
 of pseudogenes.";
 RN Mol. Biol. Evol. 2:347-358(1985).
 RN [5]
 RP SEQUENCE OF 477-551 FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=90269606; PubMed=1971802;
 RA Boccaccio C., Deschattré J., Meunier-Rotival M.;
 RT "Empty and occupied insertion site of the truncated LINE-1 repeat
 located in the mouse serum albumin-encoding gene.";
 RN Gene 88:181-186(1990).
 RN [6]
 RP SEQUENCE OF 25-44.
 RC TISSUE=Liver;
 RX MEDLINE=93162044; PubMed=1286668;
 RA Gionetti C.S., Taylor J., Tollakson S.L.;
 RT "Mouse liver protein database: a catalog of proteins detected by two-
 dimensional gel electrophoresis.";
 RN Electrophoresis 13:970-991(1992).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way

```

RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
RL "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
CC Gene 165:295-296(1996).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X84842; CAA59279.1; -.
DR PIR; JC4660; S57632.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEL 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 385 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68659 MW; 07E623CAC5F60E5F CRC64;
Query Match 81.4%; Score 57; DB 1; Length 608;
Best Local Similarity 78.6%; Pred No. 0.0078;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 TVATLRETYGEMAD 14
Db 100 TVASLRDKYGEMAD 113
RESULT 7
ALBU_CANFA
ID ALBU CANFA STANDARD; PRT; 608 AA.
AC P49822; Q77705; Q9TSZ4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)

```

```

DE Serum albumin precursor (Allergen Can f 3).
GN ALB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RA Hilger C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20148667; PubMed=10669848;
RA Pandjaitan B., Swoboda I., Brandejsky-Pichler F., Rumpold H.,
RA Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog
RT albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285(2000).
RN [3]
RP SEQUENCE OF 25-48.
RX MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W., Sarkar B.;
RT "Isolation, amino acid sequence and copper(II)-binding properties of
RT peptide (1-24) of dog serum albumin.";
RL J. Biol. Chem. 249:5872-5877(1974).
RN [4]
RP SEQUENCE OF 25-38.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
RN [5]
RP SEQUENCE OF 215-478 FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=94201492; PubMed=7512102;
RA Spitzauer S., Schweiger C., Speir W.R., Pandjaitan B., Valent P.,
RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
RT "Molecular characterization of dog albumin as a cross-reactive
RT allergen.";
RL J. Allergy Clin. Immunol. 93:614-627(1994).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ133489; CAB64867.1; -.
DR EMBL; Y17737; CAA76841.1; -.
DR EMBL; S72946; AAB30434.1; -.
DR HSSP; P02768; 1E7B.
DR HSC-2DPAGE; P49822; DOG.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.

```

RP SEQUENCE FROM N.A., AND VARIANT THR-214.
RA Wu H.T., Huang M.C.;
RT "The complete cDNA sequence of bovine serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-32.
RX MEDLINE=80024278; PubMed=488109;
RA McGillivray R.T.A., Chung D.W., Davie E.W.;
RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-terminal sequence of prealbumin.";
RL Eur. J. Biochem. 98:477-485(1979).
RN [6]
RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
RA Brown J.R.;
RT "Structure of bovine serum albumin.";
RL Fed. Proc. 34:591-591(1975).
RN [7]
RP REVISIONS TO 190-195.
RA Brown J.R.;
RT Submitted (APR-1975) to the PIR data bank.
RN [8]
RP SEQUENCE OF 402-433.
RX MEDLINE=8203364; PubMed=7283978;
RA Reed R.G., Putnam F.W., Peters T. Jr.;
RT "Sequence of residues 400-403 of bovine serum albumin.";
RL Biochem. J. 191:867-868(1980).
RN [9]
RP SEQUENCE OF 19-28.
RX MEDLINE=77134075; PubMed=843354;
RA Patterson J.E., Geller D.M.;
RT "Bovine microsomal albumin: amino terminal sequence of bovine prealbumin.";
RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
RN [10]
RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.
RX MEDLINE=91083649; PubMed=2260975;
RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;
RT "Rapid confirmation and revision of the primary structure of bovine serum albumin by EMS and Frit-FAB LC/MS.";
RL Biochem. Biophys. Res. Commun. 173:639-646(1990).
RN [11]
RP SEQUENCE OF 25-41.
RX MEDLINE=88267456; PubMed=3389500;
RA Haen J.C., Lin F., Tam M.F.;
RT "Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel: a preparative method for isolating proteins for N-terminal microsequencing.";
RL Anal. Biochem. 170:1-8(1988).
RN [12]
RP SEQUENCE OF 437-451.
RA Vilbois F.;
RT Submitted (AUG-1998) to Swiss-Prot.
RN [13]
RP DISULFIDE BONDS.
RA Brown J.R.;
RT "Structure of serum albumin: disulfide bridges.";
RL Fed. Proc. 33:1389-1389(1974).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- ALLERGEN: Causes an allergic reaction in human.
CC -1- SIMILARITY: Belongs to the ALB/AFB/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; M73993; AAA51411.1; -
DR EMBL; X58989; CA41735.1; -
DR EMBL; Y17769; CA76847.1; -
DR EMBL; AF542068; AANI7824.1; -
DR HSSP; P02768; 1E7B.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
KW Polymorphism.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
FT VARIANT 214 214
FT CONFLICT 302 302
FT CONFLICT 304 305
FT CONFLICT 324 324
FT CONFLICT 394 395
FT CONFLICT 437 437
FT CONFLICT 493 494
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;
A -> T.
C -> K (IN REF. 6).
KP -> PC (IN REF. 6).
N -> D (IN REF. 6).
ST -> TS (IN REF. 6).
K -> R (IN REF. 12).
SE -> ES (IN REF. 6).
Query Match 82.9%; Score 58; DB 1; Length 607;
Best Local Similarity 84.6%; Pred.No. 0.005;
Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 VATLRETYGEMAD 14
Db 101 VASLRETYGDMAD 113
RESULT 6
ID ALBU FELCA STANDARD; PRT; 608 AA.
AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN ALB.
OS Felis silvestris catus (Cat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96194824; PubMed=8647469;

```

FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 147 192 BY SIMILARITY.
FT DISULFID 191 200 BY SIMILARITY.
FT DISULFID 223 269 BY SIMILARITY.
FT DISULFID 268 276 BY SIMILARITY.
FT DISULFID 288 302 BY SIMILARITY.
FT DISULFID 301 312 BY SIMILARITY.
FT DISULFID 339 384 BY SIMILARITY.
FT DISULFID 415 461 BY SIMILARITY.
FT DISULFID 460 471 BY SIMILARITY.
FT DISULFID 484 500 BY SIMILARITY.
FT DISULFID 499 510 BY SIMILARITY.
FT DISULFID 537 582 BY SIMILARITY.
FT DISULFID 581 590 BY SIMILARITY.
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 88.6%; Score 62; DB 1; Length 607;
Best Local Similarity 92.3%; Pred. No. 0.00089;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VATLRETYGEMAD 14
Db 101 VATLRETYGDMAD 113

RESULT 4
ALBU_HORSE
ID ALBU_HORSE STANDARD; PRT; 607 AA.
AC P35747;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Equ c 3).
GN ALB.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
PP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93345495; PubMed=8344282;
RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
RT "X-ray and primary structure of horse serum albumin (Equus caballus)
RT at 0.27-nm resolution.";
RL Eur. J. Biochem. 215:205-212(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
-----
DR EMBL; X74045; CAA52194.1; .

```

```

DR PIR; S34053; ABHOS.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; Transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 211 396 ALBUMIN 1.
FT DOMAIN 403 594 ALBUMIN 2.
FT METAL 27 27 ALBUMIN 3.
FT DISULFID 77 86 COPPER (BY SIMILARITY).
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 68598 MW; 256F6B830A1B90C5 CRC64;

Query Match 87.1%; Score 61; DB 1; Length 607;
Best Local Similarity 85.7%; Pred. No. 0.0014;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
Db 100 TVATLRETYGELAD 113

RESULT 5
ALBU_BOVIN
ID ALBU_BOVIN STANDARD; PRT; 607 AA.
AC P02769; O02787;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Bos d 6).
GN ALB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
PP SEQUENCE FROM N.A.
RA Holowachuk E.W., Stoltzenberg J.K., Reed R.G., Peters T. Jr.;
RT Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
PP SEQUENCE FROM N.A., AND VARIANT THR-214.
RC TISSUE=Liver;
RA Barry T., Power S., Cannon F.;
RN [3]
PP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [4]

```


FT METAL 19 19 COPPER (BY SIMILARITY).
 FT BINDING 256 256 BILIRUBIN (POTENTIAL).
 FT DISULFID 69 78 BY SIMILARITY.
 FT DISULFID 91 107 BY SIMILARITY.
 FT DISULFID 106 117 BY SIMILARITY.
 FT DISULFID 140 185 BY SIMILARITY.
 FT DISULFID 184 193 BY SIMILARITY.
 FT DISULFID 216 262 BY SIMILARITY.
 FT DISULFID 261 269 BY SIMILARITY.
 FT DISULFID 281 295 BY SIMILARITY.
 FT DISULFID 294 305 BY SIMILARITY.
 FT DISULFID 332 377 BY SIMILARITY.
 FT DISULFID 376 385 BY SIMILARITY.
 FT DISULFID 408 454 BY SIMILARITY.
 FT DISULFID 453 464 BY SIMILARITY.
 FT DISULFID 477 493 BY SIMILARITY.
 FT DISULFID 492 503 BY SIMILARITY.
 FT DISULFID 530 575 BY SIMILARITY.
 FT DISULFID 574 583 BY SIMILARITY.
 SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;
 Query Match 100.0%; Score 70; DB 1; Length 600;
 Best Local Similarity 100.0%; Pred. No. 2.8e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TVATLRETYGEMAD 14
 Db 92 TVATLRETYGEMAD 105
 RESULT 2
 ALBU HUMAN STANDARD; PRT; 609 AA.
 ID ALBU HUMAN STANDARD; PRT; 609 AA.
 AC P02768; O95574; Q13140; Q9P117; Q9UHS3; Q9UJZ0;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_taxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. PubMed=3009475;
 RX MEDLINE=86196112; PubMed=3009475;
 RA Minghetti P.P., Ruffner D.B., Kuang W.J., Dennison O.E., Hawkins J.W.,
 RA Beattie W.G., Dugaiczky A.;
 RT "Molecular structure of the human albumin gene is revealed by
 RT nucleotide sequence within q11-22 of chromosome 4.";
 RL J. Biol. Chem. 261:6747-6757(1986).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
 RX MEDLINE=82081982; PubMed=6171778;
 RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
 RA Najarian R.C., Seeburg P.H., Wion K.L.;
 RT "The sequence of human serum albumin cDNA and its expression in E.
 RT coli";
 RL Nucleic Acids Res. 9:6103-6114(1981).
 RN [3]
 RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
 RX MEDLINE=82105994; PubMed=6275391;
 RA Dugaiczky A., Law S.W., Dennison O.E.;
 RT "Nucleotide sequence and the encoded amino acids of human serum
 RT albumin mRNA";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).

RC TISSUE=Fetal liver;
 RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
 RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
 RT "Functional prediction of the coding sequences of 121 new genes
 RT deduced by analysis of cDNA clones from human fetal liver."
 RN Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
 RA Huang M.C., Wu H.T.;
 RT "The cDNA sequences of human serum albumin."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver, and Skeletal muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haeh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stachenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Young A.C., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SEQUENCE OF 25-609.
 RX MEDLINE=76187907; PubMed=1225573;
 RA Meloun B., Moravsek L., Kostka V.;
 RT "Complete amino acid sequence of human serum albumin."
 RL FEBS Lett. 58:134-137(1975).
 RN [9]
 RP SEQUENCE OF 25-609.
 RA Brown J.R., Shockley P., Behrens P.Q.;
 RL (In) Bing D.H. (eds.);
 RL The chemistry and physiology of the human plasma proteins, pp.23-40,
 RL Pergamon Press, New York (1979).
 RN [10]
 RP SEQUENCE OF 1-455 FROM N.A.
 RC TISSUE=Liver;
 RA Venaya J., Parrilla R., Ayuso M.S.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE=86140099; PubMed=2419329;
 RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
 RT "The human albumin gene. Characterization of the 5' and 3' flanking
 RT regions and the polymorphic gene transcripts."
 RL J. Biol. Chem. 261:3244-3251(1986).
 RN [12]
 RP SEQUENCE OF 222-229.
 RX MEDLINE=76257808; PubMed=955075;
 RA Walker J.E.;
 RT "Lysine residue 199 of human serum albumin is modified by
 RT acetylsalicylic acid."
 RL FEBS Lett. 56:173-175(1976).
 RN [13]
 RP SEQUENCE OF 25-44 AND 480-499.
 RC TISSUE=Heart;
 RX MEDLINE=95203287; PubMed=7895732;
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 RT 1994.";

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 1.02124 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_76_89

Perfect score: 70
Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	600	1	ALBU_MACMU
2	70	100.0	609	1	ALBU_HUMAN
3	62	88.6	609	1	ALBU_SHEEP
4	61	87.1	607	1	ALBU_HORSE
5	58	82.9	607	1	ALBU_BOVIN
6	57	81.4	608	1	ALBU_FELCA
7	54	77.1	608	1	ALBU_CANFA
8	46	65.7	608	1	ALBU_MOUSE
9	43	61.4	608	1	ALBU_RABIT
10	42	60.0	256	1	PHMX_MOUSE
11	42	60.0	605	1	ALBU_PIG
12	42	60.0	608	1	ALBU_RAT
13	41	58.6	615	1	ALBU_CHICK
14	41	58.6	1034	1	GCSP_YEAST
15	40	57.1	409	1	ODCO_HABIN
16	39	57.1	563	1	PK66_PLAKU
17	39	55.7	419	1	LWA_HYDEC
18	38	54.3	420	1	YTF3_RHIT
19	38	54.3	780	1	PPSA_DEIRA
20	37	52.9	245	1	NDDB_CAVPO
21	37	52.9	312	1	OGGI_MERTH
22	37	52.9	488	1	IMDH_BACSU
23	37	52.9	759	1	TOP3_CABEL
24	37	52.9	1261	1	APAF_BRARE
25	36	51.4	283	1	Y43F_MYCPN
26	36	51.4	352	1	RFC5_NEUCR
27	36	51.4	433	1	APB_CHICK
28	36	51.4	562	1	EZRA_BACSU
29	36	51.4	614	1	DNAX_LACAC
30	36	51.4	736	1	VP4_ROTFC
31	36	51.4	1737	1	CO4_FATP
32	36	51.4	3988	1	POLG_BVDVN
33	35.5	50.7	3164	1	TEGU_HSV11

RESULT 1

ID	ALBU_MACMU	STANDARD;	PRT;	600 AA.
DT	01-NOV-1997	(Rel. 35, Created)		
DT	01-NOV-1997	(Rel. 35, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Serum albumin precursor (Fragment).			
GN	ALB.			
OS	Macaca mulatta (Rhesus macaque).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;			
OC	Cercopithecoidea; Macaca.			
OX	NCBI_TaxID=9544;			
RN	[1]_SEQUENCE FROM N.A.			
RP	MEDLINE=93211971; PubMed=8460152;			
RX	Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,			
RA	Dwulet J., Putnam F.W.;			
RT	"cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding,"			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).			
CC	FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.			
CC	SUBCELLULAR LOCATION: Secreted.			
CC	TISSUE SPECIFICITY: Plasma.			
CC	SIMILARITY: Belongs to the ALB/APP/VDB family.			
CC	SIMILARITY: Contains 3 albumin domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; M90463; AAA36906.1; -			
DR	PIR; A47391; A47391.			
DR	HSP; P02768; 1E7B.			
DR	InterPro: IPR000264; Serum albumin.			
DR	Pfam; PF00273; transport_prot; 3.			
DR	PRINTS; PR00802; SERUMALBUMIN.			
DR	ProDom; PD002486; Serum albumin; 1.			
DR	SMART; SM00103; ALBUMIN; 3.			
DR	PROSITE; PS00212; ALBUMIN; 3.			
KW	Metal-binding; Lipid-binding; Repeat; Signal; Copper.			
FT	NON_TER 1			
FT	SIGNAL 1..10			
FT	PROPEP 11..16			
FT	CHAIN 17..600			
FT	DOMAIN 17..197			
FT	DOMAIN 204..389			
FT	DOMAIN 396..587			

ALIGNMENTS

A/Cross-references: EMBL:Z47815; NID:G642280; PIDN:CAA87810.1; PID:G642281; MIPS:YMR189C
R/Sinclair, D.A.; Dawes, I.W.
submitted to the EMBL Data Library, February 1995
A/Description: Cloning and expression of glycine decarboxylase from *Saccharomyces cerevisiae*
A/Reference number: S59809
A/Accession: S59810
A/Molecule type: DNA
A/Residues: 1-1034 <SIN>
A/Cross-references: EMBL:U20641; NID:G676869; PIDN:AA818933.1; PID:G676871
R/Sinclair, D.A.; Hong, S.P.; Dawes, I.W.
Mol. Microbiol. 19, 611-623, 1996
A/Title: Specific induction by glycine of the gene for the P-subunit of glycine decarboxylase
A/Reference number: S70896; MUID:96228709; PMID:8830251
A/Accession: S70896
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-77 'V', 79-121, 'VS', 124-247, 'L', 249-381, 'T', 383-517, 'H', 519-1034 <SIW>
A/Cross-references: EMBL:U20641
C/Genetics:
A/Gene: SGD:GCV2; GCV2; GSD2
A/Cross-references: SGD:S0004801
A/Map position: 13R
C/Keywords: phosphoprotein; pyridoxal phosphate; transferase
F/773/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 58.8%; Score 41; DB 2; Length 1034;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TVATLRETYGEM 12
||| | :||| :
DB 1006 TVARLDDTYGDM 1017

RESULT 15
A95074
ABC transporter, ATP-binding protein SP0636 [imported] - *Streptococcus pneumoniae* (strain
C/Species: *Streptococcus pneumoniae*
C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C/Accession: A95074
R/Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A/Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A/Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
A/Reference number: A95000; MUID:21357209; PMID:11463916
A/Accession: A95074
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-330 <KUR>
A/Cross-references: GB:AE005672; PIDN:AAK74786.1; PID:G14972111; GSPDB:GN00164; TIGR:SP4
A/Experimental source: strain TIGR4
C/Genetics:
A/Gene: SP0636
C/Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 57.1%; Score 40; DB 2; Length 330;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVATLRETYGEM 12
||| | :||| :
DB 240 TVSQLKTFGKM 251

Search completed: April 19, 2004, 12:02:23
Job time : 2.70637 secs

A;Residues: 1-5 <RES>
A;Cross-references: GB:M16825; NID:9202828; PIDN:AAA40712.1; PID:G554412
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-608/Product: serum albumin #status experimental <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;77/Binding site: copper (His) #status experimental
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-395,384-393,4
Query Match 60.0%; Score 42; DB 1; Length 608;
Best Local Similarity 53.8%; Pred. No. 13;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 VATURETYGEMAD 14
Db 101 IPKLRDNYGELAD 113
RESULT 11
H89009
Protein R08F11.4 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C;Accession: H89009
F;Anonymous; The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: H89009
A;Status: preliminary
A;Molecule type: DNA
A;Cross-references: GB:chr_V; PIDN:AA854246.1; PID:G2088829; GSPDB:GN00023; CESP:R08F11.
A;Residues: 1-354 <STO>
A;Note: weak similarity to methyltransferases
C;Genetics:
A;Gene: R08F11.4
A;Map position: 5
Query Match 58.6%; Score 41; DB 2; Length 354;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 VATURETYGEMA 13
Db 280 VFTDRTYGRMA 291
RESULT 12
ABCS
serum albumin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C;Accession: S15571; A05078; A13451
R;Casasady, A. I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A;Reference number: S15571
A;Accession: S15571
A;Molecule type: mRNA
A;Residues: 1-615 <CAS>
R;Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, R.N.; Lau, P.C.K.; Desley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A;Title: The 5' noncoding and flanking regions of the avian very low density apolipoprob
A;Reference number: A05078; MUID:83161037; PMID:6187737
A;Accession: A05078
A;Molecule type: DNA
A;Residues: 1-28 <HAC>

A;Cross-references: GB:V00381; NID:G63038; PIDN:CRA23680.1; PID:G63039
R;Rosen, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A;Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A;Reference number: A13451; MUID:78019943; PMID:9111327
A;Accession: A13451
A;Molecule type: protein
A;Residues: 19-23, 'M', 25-30 <ROS>
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, m
ones (weak bonds with these hormones promote their transfer across the membranes), th
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-26/Domain: propeptide #status predicted <PRO>
F;27-613/Product: serum albumin #status predicted <MAT>
F;225-398/Domain: serum albumin repeat homology <SA1>
F;417-596/Domain: serum albumin repeat homology <SA2>
F;417-596/Domain: serum albumin repeat homology <SA3>
F;30/Binding site: copper (His) #status predicted
F;80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-39
Query Match 58.6%; Score 41; DB 1; Length 615;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 VATURETYGEMAD 14
Db 104 VEKLRDSYGEMAD 116
RESULT 13
G87415
glycyl-tRNA synthetase, beta subunit [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: G87415
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, C
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolc
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.N
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87415
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-692 <STO>
A;Cross-references: GB:AE005673; NID:gl3422689; PIDN:AAK23323.1; GSPDB:GN00148
C;Genetics:
A;Gene: CCI342
C;Superfamily: glycine-tRNA ligase beta chain
Query Match 58.6%; Score 41; DB 2; Length 692;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 TVATLRETYGEMAD 14
Db 673 TLATVRDAMQGVAD 696
RESULT 14
S50917
aminomethyltransferase (EC 2.1.2.10) - yeast (Saccharomyces cerevisiae)
N;Alternate names: glycyl cleavage system protein P; protein YMR189C
C;Species: Saccharomyces cerevisiae
C;Date: 10-Feb-1995 #sequence_revision 12-May-1995 #text_change 19-Apr-2002
C;Accession: S50917; S59810; S70896
R;Pearson, D.; Bowman, S.
submitted to the EMBL Data Library, January 1995
A;Reference number: S50917
A;Accession: S50917
A;Molecule type: DNA
A;Residues: 1-1034 <PEA>

F:123-236/Domain: serum albumin repeat homology <SA2>
 F:315-453/Domain: serum albumin repeat homology (fragment) <SA3>
 Query Match 65.7%; Score 46; DB 2; Length 453;
 Best Local Similarity 63.5%; Pred. No. 1.7;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VATLRETYGEMAD 14
 : ||||| |||
 Db 3 IPNLRENYGELAD 15

RESULT 8

AH2433
 hypothetical protein all15024 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AH2433
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759940
 A:Accession: AH2433
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-252 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAE76723.1; PID:gl1734162; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all15024

Query Match 60.08; Score 42; DB 2; Length 252;
 Best Local Similarity 64.3%; Pred. No. 4.9;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
 : ||||| |||
 Db 176 TIATLRSTAEHWMD 169

RESULT 9

ABPGS
 serum albumin precursor - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S01382; A61006
 R:Wainstock, J.; Baldwin, G.S.
 Nucleic Acids Res. 16, 5045, 1988
 A:Title: Nucleotide sequence of porcine liver albumin.
 A:Reference number: S01382; MUID:89016582; PMID:3174440
 A:Accession: S01382
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-605 <WEI>
 A:Cross-references: EMBL:X12422; NID:g1875; PIDN:CAA30970.1; PID:g933798
 R:Limeback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
 J. Bone Miner. Res. 4, 235-241, 1989
 A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of mineral J. Bone Miner. Res. 4, 235-241, 1989
 A:Reference number: A61006; MUID:89259769; PMID:2728927
 A:Accession: A61006
 A:Molecule type: protein
 A:Residues: 23-51, 'X', 53-54, 'XXXGY', 146, 'E', 148, 'E', 150-151, 'XVN', 155 <LIM>
 A:Experimental source: dental enamel
 A:Note: albumin and other serum proteins are also found in bone
 C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
 F:17-22/Domain: signal sequence (fragment) #status predicted <PRO>
 F:23-605/Product: serum albumin #status predicted <MAT>

F:27-199/Domain: serum albumin repeat homology <SA1>
 F:218-391/Domain: serum albumin repeat homology <SA2>
 F:410-589/Domain: serum albumin repeat homology <SA3>
 F:410-589/Domain: serum albumin repeat homology <SA3>
 F:75-84, 97-113, 112-123, 145-190, 189-198, 221-267, 266-274, 286-300, 299-310, 337-382, 381-390,
 F:261/Binding site: bilirubin (lys) #status predicted

Query Match 60.0%; Score 42; DB 1; Length 605;
 Best Local Similarity 53.8%; Pred. No. 13;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 VATLRETYGEMAD 14
 : ||||| |||
 Db 99 IPSLREHYGLAD 111

RESULT 10

ASRTS
 serum albumin precursor - rat
 N:Alternate names: preproalbumin
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 22-Jun-1999
 C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
 R:Sargent, T.D.; Yang, M.; Bonner, J.
 Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
 A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
 A:Reference number: A93872; MUID:81223722; PMID:7017712
 A:Accession: A93872
 A:Molecule type: mRNA
 A:Residues: 1-608 <SAR>
 A:Cross-references: GB:V01222; GB:J00698; NID:g55627; PIDN:CAA24532.1; PID:g55628
 R:Strauss, A.W.; Bennett, C.D.; Donchue, A.M.; Rodkey, J.A.; Alberts, A.W.
 J. Biol. Chem. 252, 6846-6855, 1977
 A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Anal.
 A:Reference number: A92211; MUID:77249657; PMID:893447
 A:Note: cleavages during protein maturation
 A:Accession: A92211
 A:Molecule type: protein
 A:Residues: 1-38 <STR>
 R:Isemura, S.; Ikenaka, T.
 J. Biochem. 83, 35-48, 1978
 A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
 A:Reference number: A91946; MUID:78109429; PMID:564345
 A:Accession: A91946
 A:Molecule type: protein
 A:Residues: 25-222 <ISI>
 R:Isemura, S.; Ikenaka, T.
 J. Biochem. 79, 1183-1196, 1976
 A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino
 A:Reference number: A91940; MUID:76260153; PMID:956149
 A:Accession: A91940
 A:Molecule type: protein
 A:Residues: 223-288; 572-608 <IS2>
 A:Note: 262-Leu was also found
 R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
 Cancer Res. 38, 3483-3486, 1978
 A:Title: Copper(II)-binding ability of human alpha-fetoprotein.
 A:Reference number: A30758; MUID:79001617; PMID:80265
 A:Contents: annotation; copper binding
 R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A:Title: Structures of histamine-releasing peptides formed by the action of acid protease
 A:Reference number: A45800; MUID:89341406; PMID:2474609
 A:Accession: C45800
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 166-173 <CAR>
 R:Heard, J.
 Mol. Cell. Biol. 7, 2425-2434, 1987

A:Title: Determinants of rat albumin promoter tissue specificity analyzed by an improve
 A:Reference number: I57621; MUID:87286876; PMID:3475566
 A:Accession: I57621
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A>Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A/Reference number: A91258; MUID:80024278; PMID:488109
A/Accession: A91258
A/Molecule type: protein
A/Residues: 1-32 <MAG>
R/Hsieh, J.C.; Lin, F.P.; Tam, M.F.
Anal. Biochem. 170, 1-8, 1988
A>Title: Electrophoretic transfer of protein from an analytical isoelectrofocusing gel.
A/Reference number: A60808; MUID:88267456; PMID:3389500
A/Accession: B60808
A/Molecule type: protein
A/Residues: 25-41 <HSI>
R/Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A>Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albumin.
A/Reference number: S10780; MUID:90336641; PMID:2379503
A/Accession: S10780
A/Molecule type: protein
A/Residues: 25-41, 'H', 43-57, 59-64 <STR>
R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A>Title: Structures of histamine-releasing peptides formed by the action of acid proteases.
A/Reference number: A45800; MUID:89341406; PMID:2474609
A/Accession: A45800
A/Molecule type: protein
A/Residues: 163-172 <CAR>
R/Carraway, R.E.; Mitra, S.P.; Cochran, D.E.
J. Biol. Chem. 262, 5968-5973, 1987
A>Title: Structure of a biologically active neurotensin-related peptide obtained from peptides.
A/Reference number: A26693; MUID:87194805; PMID:2437111
A/Accession: A26693
A/Molecule type: protein
A/Residues: 163-172, 'L', <CA2>
R/Reed, R.G.; Putnam, F.W.; Peters Jr., T.
Biochem. J. 191, 867-868, 1980
A>Title: Sequence of residues 400-403 of bovine serum albumin.
A/Reference number: A90309; MUID:82023364; PMID:7283978
A/Accession: A90309
A/Molecule type: protein
A/Residues: 402-433 <REE>
R/Brown, J.R.
Fed. Proc. 34, 591, 1975
A>Title: Structure of bovine serum albumin.
A/Reference number: A91458
A/Accession: A91458
A/Molecule type: protein
A/Residues: 190-195 <BR2>
R/Brown, J.R.
Fed. Proc. 33, 1389, 1974
A/Reference number: A91457
A/Contents: annotation; disulfide bonds
R/Werlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A>Title: Preparation and characterization of novel substrates of insulin proteinase (EC 3.4.21.3).
A/Reference number: S55232; MUID:95031935; PMID:7945219
A/Accession: S55232
A/Status: preliminary
A/Molecule type: protein
A/Residues: 529-536; 569-572 <WER>
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: carrier protein; copper binding; duplication; plasma
F/1-18/Domain: signal sequence #status experimental <SIG>
F/19-24/Domain: propeptide #status experimental <PRO>
F/25-607/Product: serum albumin #status experimental <MPT>
F/29-201/Domain: serum albumin repeat homology <SA1>
F/220-393/Domain: serum albumin repeat homology <SA2>
F/412-591/Domain: serum albumin repeat homology <SA3>

F/27/Binding site: copper (His) #status predicted
F/77-86, 99-115, 114-125, 147-192, 191-200, 223-269, 268-276, 288-302, 301-312, 339-384, 383-392,
Query Match 82.9%; Score 58; DB 1; Length 607;
Best Local Similarity 84.6%; Pred. No. 0.015;
Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 VATLRETYGEMAD 14
DB 101 VASLRDXYGEMAD 113
RESULT 6
S57632
serum albumin precursor - cat
C/Species: Felis silvestris catus (domestic cat)
C/Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
C/Accession: JC4660; S57632
R/Hilger, C.; Grigioni, F.; Hentges, F.
Gene 169, 295-296, 1996
A/Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A/Reference number: JC4660; MUID:96194824; PMID:8647469
A/Accession: JC4660
A/Molecule type: mRNA
A/Residues: 1-608 <HI2>
A/Cross-references: EMBL:X84842; NID:G986484; PIDN:CAA59279.1; PID:G986485
A/Experimental source: liver
C/Comment: This protein is the major protein component in plasma. It functions as a mu
ein has 35 conserved cysteine residues.
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: liver; plasma
F/1-18/Domain: signal sequence #status predicted <SIG>
F/19-24/Domain: propeptide #status predicted <PR>
F/25-608/Product: serum albumin #status predicted <MAT>
F/29-202/Domain: serum albumin repeat homology <SA1>
F/221-394/Domain: serum albumin repeat homology <SA2>
F/413-592/Domain: serum albumin repeat homology <SA3>
Query Match 81.4%; Score 57; DB 2; Length 608;
Best Local Similarity 78.6%; Pred. No. 0.023;
Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 TVATLRETYGEMAD 14
DB 100 TVASLRDXYGEMAD 113
RESULT 7
A05139
serum albumin - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C/Accession: A05139; I48638
R/Minghetti, P.P.; Law, S.W.; Dugalczyk, A.
Mol. Biol. Evol. 2, 347-358, 1985
A/Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseud
A/Reference number: A93055; MUID:88216123; PMID:2452956
A/Accession: A05139
A/Molecule type: mRNA
A/Residues: 1-418 <MIN>
A/Cross-references: GB:ML6111; NID:gl91764; PIDN:AAA37190.1; PID:gl91765
R/Boccaccio, C.; Deschatrette, J.; Meunier-Rottival, M.
Gene 88, 181-186, 1990
A/Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in th
A/Reference number: I48638; MUID:90269606; PMID:1971802
A/Accession: I48638
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 379-453 <BOC>
A/Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: carrier protein; duplication; metal binding; plasma
F/1-104/Domain: serum albumin repeat homology (fragment) <SA1>

A;Contents: annotation; binding sites
 R;Harper, M.E.; Dugaiczak, A.
 Am. J. Hum. Genet. 35, 565-572, 1983
 A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein genes
 A;Reference number: A90028; MUID:83279982; PMID:6192711
 A;Contents: annotation; gene position
 R;Walker, J.E.
 FEBS Lett. 66, 173-175, 1976
 A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid.
 A;Reference number: A46755; MUID:76257808; PMID:955075
 A;Contents: annotation
 R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
 FEBS Lett. 298, 266-268, 1992
 A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosphate
 A;Reference number: A56294; MUID:92183881; PMID:1544460
 A;Contents: annotation
 A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in p
 atase activity
 C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
 in liver, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
 C;Comment: A large number of variants of human serum albumin have been described.
 C;Genetics:
 A;Gene: GDB:ALB
 A;Cross-references: GDB:118990; OMIM:103600
 A;Map Position: q41-q41.3
 C;Superfamily: serum albumin; serum albumin repeat homology
 C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyridox
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-24/Domain: propeptide #status experimental <PRO>
 F;25-609/Product: serum albumin #status experimental <MPt>
 F;28-202/Domain: serum albumin repeat homology <SA1>
 F;166-174/Product: kinetensin #status experimental <KTP>
 F;221-394/Domain: serum albumin repeat homology <SA2>
 F;413-592/Domain: serum albumin repeat homology <SA3>
 F;27/Binding site: copper (His) #status predicted
 F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4
 F;214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 70; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 9.1e-05; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
 |||||
 DB 100 TVATLRETYGEMAD 113

RESULT 3
 ABSHS
 serum albumin precursor - sheep
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C;Accession: S06936
 R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
 Nucleic Acids Res. 17, 10495, 1989
 A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
 A;Reference number: S06936; MUID:90098888; PMID:2602160
 A;Accession: S06936
 A;Molecule type: mRNA
 A;Residues: 1-607 <PRO>
 A;Cross-references: EMBL:X17055; NID:G1386; PIDN:CAA34903.1; PID:G1387
 C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
 steroid hormones (weak bonds with these hormones promote their transfer across the membra
 C;Superfamily: serum albumin; serum albumin repeat homology
 C;Keywords: carrier protein; duplication; metal binding; plasma
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-24/Domain: propeptide #status predicted <PRO>
 F;25-607/Product: serum albumin #status predicted <MAT>
 F;29-201/Domain: serum albumin repeat homology <SA1>
 F;220-393/Domain: serum albumin repeat homology <SA2>
 F;412-591/Domain: serum albumin repeat homology <SA3>
 F;27/Binding site: copper (His) #status predicted

F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,
 F;263/Binding site: bilirubin (Lys) #status predicted

Query Match 88.6%; Score 62; DB 1; Length 607;
 Best Local Similarity 92.3%; Pred. No. 0.0027;
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VATLRETYGEMAD 14
 |||||
 DB 101 VATLRETYGEMAD 113

RESULT 4
 ABHOS
 serum albumin precursor - horse
 C;Species: Equus caballus (domestic horse)
 C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C;Accession: S34053
 R;Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
 Eur. J. Biochem. 215, 205-212, 1993
 A;Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
 A;Reference number: S34053; MUID:93345495; PMID:8344282
 A;Accession: S34053
 A;Molecule type: mRNA
 A;Residues: 1-607 <HOA>
 A;Cross-references: GB:X74045; NID:G399671; PIDN:CAA52194.1; PID:G399672
 C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
 steroid hormones (weak bonds with these hormones promote their transfer across the membe
 C;Superfamily: serum albumin; serum albumin repeat homology
 C;Keywords: carrier protein; duplication; metal binding; plasma
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-24/Domain: propeptide #status predicted <PRO>
 F;25-607/Product: serum albumin #status predicted <MAT>
 F;29-201/Domain: serum albumin repeat homology <SA1>
 F;220-393/Domain: serum albumin repeat homology <SA2>
 F;412-591/Domain: serum albumin repeat homology <SA3>
 F;27/Binding site: copper (His) #status predicted
 F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,
 F;263/Binding site: bilirubin (Lys) #status predicted

Query Match 87.1%; Score 61; DB 1; Length 607;
 Best Local Similarity 85.7%; Pred. No. 0.0041;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
 |||||
 DB 100 TVATLRETYGEMAD 113

RESULT 5
 ABHOS
 serum albumin precursor [validated] - bovine
 N;Alternate names: 67K protein; preproalbumin
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
 C;Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A;
 R;Holowachuk, E.W.; Stoltzenberg, J.K.; Reed, R.G.; Peters Jr., T.
 submitted to the EMBL Data Library, August 1991
 A;Description: Bovine serum albumin: cDNA sequence and expression.
 A;Reference number: A38885
 A;Accession: A38885
 A;Molecule type: mRNA
 A;Residues: 1-607 <HOA>
 A;Cross-references: EMBL:M73215
 R;Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
 Biochem. Biophys. Res. Commun. 173, 639-646, 1990
 A;Title: Rapid confirmation and revision of the primary structure of bovine serum album
 A;Reference number: A36401; MUID:91083649; PMID:2260975
 A;Accession: A36401
 A;Molecule type: protein
 A;Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>
 R;Macgillivray, R.T.A.; Chung, D.W.; Davie, E.W.
 Eur. J. Biochem. 98, 477-485, 1979

A/Accession: A93936
A/Molecule type: mRNA
A/Residues: 1-120, 'G', 122-609 <DUG>
A/Cross-references: EMBL:V00494; NID:928589; PIDN:CAA23753.1; PID:928590
R/Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A/Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and A/Reference number: I39427; MUID:86140099; PMID:2419329
A/Accession: I39427
A/Status: translation not shown
A/Molecule type: DNA
A/Residues: 1-26 <URA>
A/Cross-references: GB:M13075; NID:G178330; PIDN:AA51688.1; PID:9553173
R/Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A/Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family. A/Reference number: I59286; MUID:94181575; PMID:8134387
A/Accession: I59286
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 282-290, 'KSRFDLO' <WAT>
A/Cross-references: GB:S69192; NID:G546032; PIDN:RAB30282.1; PID:9546033
A/Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R/Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam, Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A/Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl- A/Reference number: I59313; MUID:94294404; PMID:8022807
A/Accession: I59313
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 589-590, 'ALPRVRNLLQVLP' <MAD>
A/Cross-references: GB:S70759; NID:G547231; PIDN:RAB31177.1; PID:9547232
A/Note: this frame-shift variant is designated albumin Bazzano; four additional variants submitted to the EMBL Data Library, March 1995
R/Menaya, J.; Parrilla, R.; Ayuso, M.S.
A/Reference number: G08292
A/Accession: G01747
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-120, 'G', 122-455 <MEN>
A/Cross-references: EMBL:U25966; NID:G763428; PIDN:AAA64922.1; PID:9763431
R/Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A/Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2 A/Reference number: S55314; MUID:95275251; PMID:7755581
A/Accession: S55314
A/Molecule type: protein
A/Residues: 19-27 <LED>
R/Meloun, B.; Moravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A/Title: Complete amino acid sequence of human serum albumin.
A/Reference number: A91420; MUID:76187907; PMID:1225573
A/Accession: A91420
A/Molecule type: protein
A/Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395- R/Koehn, U.; Spittler, G.; Tripiet, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A/Title: Isolation and structure elucidation of middle-molecular weight peptides from ur A/Reference number: S06422
A/Note: this paper is in German, with an English abstract
A/Accession: S06422
A/Molecule type: protein
A/Residues: 25-48 <ROD>
R/Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A/Title: Mass spectrometric identification of modifications to human serum albumin treat A/Reference number: S36882; MUID:93384321; PMID:8373198
A/Accession: S36882
A/Molecule type: protein
A/Residues: 45-67, 141-160, 311-337, 469-490, 570-581 <PIN>
R/Kausler, E.; Spittler, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A/Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol

A/Reference number: S17599; MUID:92126241; PMID:1772598
A/Accession: S17599
A/Molecule type: protein
A/Residues: 25-54, 1354-357, 431-447 <KAU>
A/Note: 49-Leu was also found
R/Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1694, 1989
A/Title: Structures of histamine-releasing peptides formed by the action of acid protease A/Reference number: A45800; MUID:89341406; PMID:2474609
A/Accession: A45800
A/Molecule type: protein
A/Residues: 166-173, 'L' <MOG>
R/Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; P Biochem. Biophys. Res. Commun. 136, 983-988, 1986
A/Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-ti A/Reference number: A03239; MUID:86242180; PMID:3087352
A/Accession: A03239
A/Molecule type: protein
A/Residues: 76-83, 'K', 85-106 <GAL3>
A/Note: this variant is designated albumin Torino
R/Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Temi, R.
Eur. J. Biochem. 214, 437-444, 1993
A/Title: The structural characterization and bilirubin-binding properties of albumin He A/Reference number: S33298; MUID:93292504; PMID:8513793
A/Accession: S33298
A/Molecule type: protein
A/Residues: 255-263, 'E', 265-281 <MIN1>
A/Note: this variant is designated albumin Herborn
R/Minchiotti, L.; Galliano, M.; Stoppani, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Port Biochim. Biophys. Acta 1119, 232-238, 1992
A/Title: Two allolalbumins with identical electrophoretic mobility are produced by diff A/Reference number: S21078; MUID:92190239; PMID:1347703
A/Accession: S21078
A/Molecule type: protein
A/Residues: 354-356, 'K', 358-378 <MIN2>
A/Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported R/He, X.M.; Carter, D.C.
Nature 359, 209-215, 1992
A/Title: Atomic structure and chemistry of human serum albumin.
A/Reference number: A46756; MUID:92334427; PMID:1630489
A/Contents: annotation; X-ray crystallography, 2.8 angstroms
R/Brown, J.R.; Shockey, P.; Behrens, F.Q.
In The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40 A/Reference number: A94442
A/Contents: annotation; three-dimensional structure and disulfide bonds
R/Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A/Title: Disulfide bonds in human serum albumin.
A/Reference number: A90930
A/Contents: annotation; disulfide bonds
R/Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A/Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding A/Reference number: A90299; MUID:78186630; PMID:656055
A/Contents: annotation; bilirubin-binding site
R/Peters, T.; Reed, R.G.
In Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20 A/Title: Serum albumin: conformation and active sites.
A/Reference number: A94408

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 1.70637 seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_copy_76_89

Perfect score: 70
Sequence: 1 TVATLRETYGEMAD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: Piri:*

2: Piri:*

3: Piri:*

4: Piri:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	100.0	600	2 A47391	serum albumin prec
2	70	100.0	609	1 ABHUS	serum albumin prec
3	62	88.6	607	1 ABHSH	serum albumin prec
4	61	87.1	607	1 ABHOS	serum albumin prec
5	58	82.9	607	1 ABHOS	serum albumin prec
6	57	81.4	608	2 S57632	serum albumin prec
7	46	65.7	453	2 A05139	serum albumin prec
8	42	60.0	282	2 AB2433	serum albumin mo
9	42	60.0	605	1 ABHPS	hypothetical prote
10	42	60.0	608	1 ABRTS	serum albumin prec
11	41	58.6	354	2 H89009	protein R08F11.4 [
12	41	58.6	615	1 ABCHS	serum albumin prec
13	41	58.6	692	2 G87415	glycyl-tRNA synthet
14	41	58.6	1034	2 S50917	aminomethyltransfer
15	40	57.1	330	2 A95074	ABC transporter, A
16	40	57.1	330	2 E97941	hypothetical prote
17	40	57.1	409	2 D64135	dihydrodiploamide s
18	40	57.1	563	2 A39238	66K merozoite surf
19	40	57.1	585	2 B75265	probable acid-CoA
20	39.5	56.4	202	2 T31452	probable 4-vinyl r
21	39	55.7	157	2 T22606	hypothetical prote
22	39	55.7	179	2 C86853	hypothetical prote
23	38	54.3	251	2 A80805	Vacu lipoprotein p
24	38	54.3	780	2 D75361	phosphoenolpyruvat
25	38	54.3	867	2 F81086	aminopeptidase N
26	37.5	53.6	1350	2 T42697	hypothetical prote
27	37	52.9	169	2 AE3257	shikimate kinase
28	37	52.9	228	2 D96945	response regulator
29	37	52.9	251	2 T46373	hypothetical prote

30	37	52.9	312	2 E69045	8-oxoguanine DNA g
31	37	52.9	332	2 H83924	tryptophanyl-tRNA
32	37	52.9	488	1 DEBSMP	IMP dehydrogenase
33	37	52.9	759	2 T43031	DNA topoisomerase
34	37	52.9	1172	2 AD2310	hypothetical prote
35	37	52.9	1278	2 T27925	hypothetical prote
36	36.5	52.1	378	2 T39790	hypothetical prote
37	36	51.4	69	2 D95250	hypothetical prote
38	36	51.4	89	2 A69513	probable dimethyl
39	36	51.4	145	2 C85896	hypothetical prote
40	36	51.4	157	2 B83496	hypothetical prote
41	36	51.4	205	2 AH0403	anaerobic dimethyl
42	36	51.4	209	2 G91051	hypothetical prote
43	36	51.4	256	2 D64637	amino acid ABC tra
44	36	51.4	257	2 E71877	probable amino aci
45	36	51.4	287	2 S73529	M6439 homolog 809_

ALIGNMENTS

RESULT 1

A47391 serum albumin precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A47391
R:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bili
A:Reference number: A47391; MUID:93211971; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-500 <WAT>
A:Cross-references: GB:M90463; NID:G342294; PIDN:AAA36906.1; PID:G342295
A:Experimental source: liver
A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:128281)
F:21-194/Domain: serum albumin; serum albumin repeat homology
F:213-386/Domain: serum albumin repeat homology <SA1>
F:405-584/Domain: serum albumin repeat homology <SA2>
F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 100.0%; Score 70; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 8.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TVATLRETYGEMAD 14
|||
DB 92 TVATLRETYGEMAD 105

RESULT 2

ABHUS serum albumin precursor [validated] - human
N:Alternate names: preproalbumin
N:Contains: kinetensin
C:Species: Homo sapiens (man)
C:Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text_change 17-Mar-2000
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S3
R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houck, C.M.; Najarian, R.C.; Seebu
Nucleic Acids Res. 9, 6103-6114, 1981
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia col
A:Reference number: A93743; MUID:82081882; PMID:6171778
A:Accession: A93743
A:Molecule type: mRNA
A:Residues: 1-419; K, 421-609 <LAW>
A:Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA2
R:Dugaiczky, A.; Law, S.W.; Dennison, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.
A:Reference number: A93936; MUID:82105994; PMID:6275391


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-984-186-2

```

```

Query Match      100.0%; Score 38; DB 4; Length 610;
Best Local Similarity 100.0%; Pred. No. 7.5; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      78 VADESAEN 85

```

RESULT 15

```

US-10-153-064-133
; Sequence 133, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 133
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-133

```

```

Query Match      100.0%; Score 38; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 8.1; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      120 VADESAEN 127

```

Search completed: April 19, 2004, 12:05:19
Job time : 2.22622 secs

```

; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04075
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US95-04075-3
;
; Query Match 100.0%; Score 38; DB 5; Length 609;
; Best Local Similarity 100.0%; Pred. No. 7.5;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 VADESAEN 8
; Db 78 VADESAEN 85
;
; RESULT 13
; US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
;
; Query Match 100.0%; Score 38; DB 2; Length 610;
; Best Local Similarity 100.0%; Pred. No. 7.5;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 VADESAEN 8
; Db 78 VADESAEN 85
;
; RESULT 14
; US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
;
; Application Number: PCT/FR93/00085
; Filing Date: 28-JAN-1993
; Attorney/Agent Information:
; Name: Smith Ph.D., Julie K.
; Registration Number: P-38,619
; Reference/Docket Number: ST92006-US
; Telecommunication Information:
; Telephone: (610) 454-3839
; Telefax: (610) 454-3808
; Information for Seq ID No: 2:

```

APPLICANT: Bart, Kathryn A.
APPLICANT: Brierley, Russell A.
APPLICANT: Thill, Gregory P.
APPLICANT: Tschopp, Juerg F.
TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
TITLE OF INVENTION: PICHIA PASTORIS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: U.S.A.
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,037
FILING DATE: 03-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DiGiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 91092
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 100.0%; Score 38; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 78 VADESAEN 85

RESULT 9

US-08-897-956A-2
Sequence 2, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/CPA
CURRENT APPLICATION NUMBER: US/08/897,956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 609
TYPE: PPT
ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 100.0%; Score 38; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 78 VADESAEN 85

RESULT 10
US-10-153-064-7
Sequence 7, Application US/10153064
Patent No. 6663485
GENERAL INFORMATION:
APPLICANT: Bell et al.
TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
FILE REFERENCE: PFS56
CURRENT APPLICATION NUMBER: US/10/153,064
CURRENT FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: 60/293,212
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 137
SOFTWARE: Patentin version 3.1
SEQ ID NO 7
LENGTH: 609
TYPE: PPT
ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match 100.0%; Score 38; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 78 VADESAEN 85

RESULT 11

US-09-976-594-977
Sequence 977, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 977
LENGTH: 609
TYPE: PPT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match 100.0%; Score 38; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 78 VADESAEN 85

RESULT 12

PCT-US95-04075-3
Sequence 3, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-702-572-2

Query Match          100.0%; Score 38; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 5
US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MERRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match          100.0%; Score 38; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 6
US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PP556

```

```

; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRM
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match          100.0%; Score 38; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 7
US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: A Human Serum Albumin-Like
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Angen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,619
; FILING DATE:
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match          100.0%; Score 38; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 78 VADESAEN 85

RESULT 8
US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan

```

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Region
LOCATION: 369..419
OTHER INFORMATION: /note= "Alternative C-termini of
OTHER INFORMATION: HSA(1-n)"
FEATURE:
NAME/KEY: Region
LOCATION: 1..585
OTHER INFORMATION: /note= "Amino acid sequence of
OTHER INFORMATION: natural HSA"
US-08-153-799-14
Query Match 100.0%; Score 38; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.2; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VADESAEN 8
DB 54 VADESAEN 61
RESULT 2
US-08-448-196A-3
; Sequence 3, Application US/08448196A
; Patent No. 5780594
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
; TITLE OF INVENTION: RELATED PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NASA
; STREET: MARSHALL SPACE FLIGHT CENTER
; CITY: HUNTSVILLE
; STATE: ALABAMA
; COUNTRY: USA
; ZIP: 35812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,196A
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: BROAD JR., ROBERT L.
; REGISTRATION NUMBER: 18,757
; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
; TELEPHONE: 205-544-0021
; TELEFAX: 205-544-0258
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-448-196A-3
Query Match 100.0%; Score 38; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.2; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
DB 54 VADESAEN 61
RESULT 3
US-08-984-176-1
; Sequence 1, Application US/08984176
; Patent No. 5948609
; GENERAL INFORMATION:
; APPLICANT: CARTER, DANIEL C
; APPLICANT: HO, JOSEPH X
; APPLICANT: RUKER, FLORIAN
; TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
; TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER
; FILE REFERENCE: 08/984,176
; CURRENT APPLICATION NUMBER: US/08/984,176
; CURRENT FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-984-176-1
Query Match 100.0%; Score 38; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.2; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VADESAEN 8
DB 54 VADESAEN 61
RESULT 4
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: MicroSoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 1.23622 Seconds
(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18_COPY_54_61
Perfect score: 38
Sequences: 1 VADESAEN 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	585	1	US-08-153-799-14
2	38	100.0	585	1	US-08-448-196A-3
3	38	100.0	585	2	US-08-984-176-1
4	38	100.0	585	2	US-08-702-572-2
5	38	100.0	585	3	US-08-769-746-2
6	38	100.0	585	4	US-10-153-064-5
7	38	100.0	609	1	US-08-222-619-3
8	38	100.0	609	1	US-08-433-037-4
9	38	100.0	609	4	US-08-897-956A-2
10	38	100.0	609	4	US-10-153-064-7
11	38	100.0	609	4	US-09-976-594-977
12	38	100.0	609	5	PCT-US95-04075-3
13	38	100.0	610	2	US-08-797-689-2
14	38	100.0	610	4	US-09-984-186-2
15	38	100.0	651	4	US-10-153-064-133
16	38	100.0	652	4	US-10-153-064-96
17	38	100.0	652	4	US-10-153-064-99
18	38	100.0	652	4	US-10-153-064-105
19	38	100.0	652	4	US-10-153-064-131
20	38	100.0	653	4	US-10-153-064-131
21	38	100.0	656	4	US-10-153-064-130
22	38	100.0	660	4	US-10-153-064-90
23	38	100.0	660	4	US-10-153-064-93
24	38	100.0	668	4	US-10-153-064-102
25	38	100.0	676	4	US-10-153-064-95
26	38	100.0	676	4	US-10-153-064-98
27	38	100.0	676	4	US-10-153-064-104

Sequence 127, App
Sequence 129, App
Sequence 125, App
Sequence 123, App
Sequence 92, Appl
Sequence 101, App
Sequence 2, Appli
Sequence 4, Appli
Sequence 16, Appl
Sequence 16, Appl
Sequence 3, Appli
Sequence 89, Appl
Sequence 7, Appli
Sequence 5, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4714, Ap
Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153.799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

of the colloidal osmotic pressure of blood.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Plasma.
CC -|- SIMILARITY: Belongs to the ALB/AFB/VDB family.
CC -|- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; M18350; AAA49636.1; --
DR PIR; A41682; ABXL68.
DR HSP; P02768; IZ7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 606
FT DOMAIN 25 205
FT DOMAIN 212 395
FT DOMAIN 402 593
FT METAL 30 30
FT METAL 80 89
FT DISULFID 102 118
FT DISULFID 117 128
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 383
FT DISULFID 382 391
FT DISULFID 414 460
FT DISULFID 459 470
FT DISULFID 483 499
FT DISULFID 498 509
FT DISULFID 536 581
FT DISULFID 580 589
FT DISULFID 580 589
SQ SEQUENCE 606 AA; 70499 MW; F8E72A952C3A3E8 CRC64;

Query Match 35.1%; Score 726.5; DB 1; Length 606;
Best Local Similarity 36.9%; Pred. No. 1.2e-46;
Matches 139; Conservative 74; Mismatches 155; Indels 9; Gaps 4;

Qy 3 HKSEVAHRFDLGENFKALVIAFAQVLOQCPEDHVKLVNEVTEPAKTCVADESAENC 62
Db 30 HKHGIADMYNLLTERTFGLTALVSNLQKCSLELSKLVNEINDFAKSGTNDKTPCC 89
Qy 63 DKSLLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHK--DDNPNLPLVRYEV 120
Db 90 EKPIGTLFYDKLCKADPKGVNMYNSKECCSKQDPERAQCFAHRVFEHNP-----VRPKP 144
Qy 121 DVMCTAFHDNEETLKKYLYEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180
Db 145 EETCALFKHEPDDLLSAFIHEEARNHPDLPYPAVLLLTQQYKLVHCCBEEDKCKFAE 204
Qy 181 KLDELREGKASSAKQLKCKASQKGFGERAFKAWAVARLSQRPFKAEFAEVSKLVTDLTK 240
Db 205 KKKELMKHSHSIEDKQKHFQWVNNYPERVVKALNARVSHRYPKDPFKLAHKTBTETH 264
Qy 241 VHTCCGHDLECCADDDADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA 300
Db 265 FIKDCCHGDMFECMTLELSEHTCOHKDELSTKLEKCCNPLPLELTYCIVTLENDVPA 324

Qy 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMPLYEYARRHPDYSVVLLLRLLAKTYETTLK 360
Db 325 ELSKPITEFTEDPHVCEKYAENKS-FLEISFWQ-SQETPELSQFLQSAKYEVSLLNKC 382
Qy 361 CAAADPHECYAKVDFEF 377
Db 383 CFSNPPCEYKDGADRF 399

Search completed: April 19, 2004, 11:52:45
Job time : 30.2299 secs


```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RA  Cassidy A.I., Salkild C.K., Baverstock P., Wallace J.C.;
RL  Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE OF 1-28 FROM N.A.
RX  MEDLINE=83161037; PubMed=6187737;
RA  Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.;
RT  "The 5' noncoding and flanking regions of the avian very low density
RT  apolipoprotein II and serum albumin genes. Homologies with the egg
RT  white protein genes.";
RL  J. Biol. Chem. 258:4556-4564(1983).
RN  [3]
RP  SEQUENCE OF 19-30.
RX  MEDLINE=78019943; PubMed=911327;
RA  Rosen A.W., Geller D.M.;
RT  "Chicken macroosomal albumin: amino terminal sequence of chicken
RT  proalbumin.";
RL  Biochem. Biophys. Res. Commun. 78:1060-1066(1977).
RN  [4]
RP  ALLERGENIC PROPERTIES.
RX  MEDLINE=21381307; PubMed=11488669;
RA  Quirice S., Maranon F., Umplierrez A., de las Heras M.,
RA  Fernandez-Caldas E., Sastre J.;
RT  "Chicken serum albumin (Gal d 5*) is a partially heat-labile inhalant
RT  and food allergen implicated in the bird-egg syndrome.";
RL  Allergy 56:754-762(2001).
CC  -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC  binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC  hormones, bilirubin and drugs. Its main function is the regulation
CC  of the colloidal osmotic pressure of blood.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- TISSUE SPECIFICITY: Plasma.
CC  -!- ALLERGEN: Causes an allergic reaction in human. Binds IgE.
CC  Partially heat-labile allergen that may cause both respiratory and
CC  food-allergy symptoms in patients with the bird-egg syndrome.
CC  -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC  -!- SIMILARITY: Contains 3 album domains.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X60688; CAA43098.1; -
DR  EMBL; V00381; CAA23680.1; -
DR  PIR; S15571; ABCHS.
DR  HSP; P02768; IE7B.
DR  InterPro; IPR000264; Serum albumin.
DR  Pfam; PF00273; transport prot; 3.
DR  PRINTS; PR00802; SERUMALBUMIN.
DR  ProDom; PD002486; Serum albumin; 1.
DR  SMART; SM00103; ALBUMIN_3.
DR  PROSITE; PS00012; ALBUMIN; 3.
KW  Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT  SIGNAL 1
FT  PROPEP 19
FT  CHAIN 24
FT  DOMAIN 24
FT  DOMAIN 216
FT  DOMAIN 408
FT  METAL 30
FT  METAL 30
FT  DISULFID 80
FT  DISULFID 102
FT  DISULFID 117
FT  DISULFID 152
FT  DISULFID 196
FT  DISULFID 228
FT  DISULFID 273

```

```

FT  DISULFID 293
FT  DISULFID 306
FT  DISULFID 317
FT  DISULFID 344
FT  DISULFID 388
FT  DISULFID 420
FT  DISULFID 466
FT  DISULFID 485
FT  DISULFID 489
FT  DISULFID 504
FT  DISULFID 542
FT  DISULFID 586
FT  DISULFID 595
FT  CARBOHYD 500
FT  CONFLICT 24
SQ  SEQUENCE 615 AA; 69918 MW; E59E4BBAEC066C6 CRC64;

Query Match
Best Local Similarity 47.9%; Score 1038.5; DB 1; Length 615;
Matches 185; Conservative 75; Mismatches 125; Indels 1; Gaps 1;

QY 3 HKSEVAHREKDLGEENFKALVLAFAQYLQCCPFEDHVKLVNNEVTFEAKTCVADESAENC 62
DB 30 HKSEIAHYNDLKEETFKAVAMITFAQLQRCSTEGSLKLVKDVLDLAQKCVANEDAPEC 89

QY 63 DKSHTLFGDKLCTVATLRETYGEMADCCAKQBFERNECFLOHKDDNPMLPR-LVRPEVD 121
DB 90 SKLPISILDEIQVEKLRDSYGAMADCCSKADPERNECFSLFKVSQPDFVQVQRPASD 149

QY 122 VMCTAFHDNEETFLKKVLYEIAIRHRYFYFAPELLFFAKRYKAATFECQQAADRAALLPK 181
DB 150 VICQEQDNVRSFLGHFIYSVAERHPFLYAPAILSFVDFEHALQSCCKESDVGACLDTK 209

QY 182 LDELRLDEGKASSAKQRLKCLASLOKFGERAPKAWARLSORFPKAPFAEYKLVTDLTKV 241
DB 210 EIVMREKAGSVKQVFCGLKQFGDRVFOARQLIYLSQYKPAFSEVSKFVHSIGV 269

QY 242 HTECHGDLLECADDRADLAKYICENQDSISSKLKSCCKEPLLEKSHCHAEVNDMPAD 301
DB 270 HKSCCEGDWYECMDMARMMSNLCSQODVFSGKIKDCCKEPIVERSQCIIMEAEFDEKPAD 329

QY 302 LPSLAADFVBSKDVCKNYAEAKDVFLGFLYEAHRRPDPVSVLLRLAKTYETTTLEKCC 361
DB 330 LPSLVEKYIIDKEVCKSFEGAGHDAFMAEFVYSRRHRPEFSIQLIWIARIKGYESLLEKCC 389

QY 362 AAADPHCYAKVDFEKPFLVEEPNL 387
DB 390 KTDNPAFCYANACEQLNQHIKETQDV 415

RESULT 14
ALB2_XENLA
ID ALB2_XENLA STANDARD; PRT; 607 AA.
AC P14872;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 74 kDa serum albumin precursor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
CX NCBI_TaxID=8355;
[1]
RN RP SEQUENCE OF 3-607 FROM N.A.
RX MEDLINE=89313788; PubMed=2747653;
RA Moskaitis J.E., Sargent T.D., Smith L.H. Jr., Pastori R.L.,
RA Schoenberg D.R.;
RT "Xenopus laevis serum albumin: sequence of the complementary
RT deoxyribonucleic acids encoding the 68- and 74-kilodalton peptides
RT and the regulation of albumin gene expression by thyroid hormone
RL during development.";
RN [2]. Endocrinol. 3:464-473(1989).
RP SEQUENCE OF 1-48 FROM N.A.
RX MEDLINE=88172470; PubMed=2451026;

```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Kryzyski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN SEQUENCE OF 99-516 FROM N.A.
RP MEDLINE=88218125; PubMed=2452956;
RX Mirghetti P.P., Law S.W., Dugaiczky A.,
RA "The rate of molecular evolution of alpha-fetoprotein approaches that
RT of pseudogenes";
RL Mol. Biol. Evol. 2:347-358(1985).
RN [5]
RN SEQUENCE OF 477-551 FROM N.A.
RP STRAIN=BALB/c;
RX MEDLINE=90269606; PubMed=1971802;
RA Boccaccio C., Deschatrette J., Meunier-Rotival M.,
RT "Empty and occupied insertion site of the truncated LINE-1 repeat
RT located in the mouse serum albumin-encoding gene";
RL Gene 88:181-186(1990).
RN [6]
RN SEQUENCE OF 25-44.
RP TISSUE=Liver;
RX MEDLINE=93162044; PubMed=1286668;
RA Giometti C.S., Taylor J., Tollakson S.L.,
RT "Mouse liver protein database: a catalog of proteins detected by two-
RT dimensional gel electrophoresis";
RL Electrophoresis 13:970-991(1992).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL; AJ011413; CAA09617.1; -
DR EMBL; AK010025; BAB26650.1; -
DR EMBL; BC049971; AAH49971.1; -
DR EMBL; M16111; AAA37190.1; -
DR EMBL; X13060; CAA31458.1; -
DR PIR; A05339; A05139.
DR HSPF; P02768; I87B.
DR SWISS-2DPAGE; P07724; MOUSE.
DR MGD; MGI:87991; Alb1.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SMC0103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR METAL-binding; lipid-binding; Repeat; Signal; Copper.
KW SIGNAL 18
FT PROPEP 19 24
FT CHAIN 25 608
FT SERUM ALBUMIN.

FT	DOMAIN	25	205	ALBUMIN 1.
FT	DOMAIN	212	397	ALBUMIN 2.
FT	DOMAIN	404	595	ALBUMIN 3.
FT	METAL	27	27	COPPER.
FT	DISULFID	77	86	BY SIMILARITY.
FT	DISULFID	99	115	BY SIMILARITY.
FT	DISULFID	114	125	BY SIMILARITY.
FT	DISULFID	148	193	BY SIMILARITY.
FT	DISULFID	192	201	BY SIMILARITY.
FT	DISULFID	224	270	BY SIMILARITY.
FT	DISULFID	269	277	BY SIMILARITY.
FT	DISULFID	289	303	BY SIMILARITY.
FT	DISULFID	302	313	BY SIMILARITY.
FT	DISULFID	340	385	BY SIMILARITY.
FT	DISULFID	384	393	BY SIMILARITY.
FT	DISULFID	416	462	BY SIMILARITY.
FT	DISULFID	461	472	BY SIMILARITY.
FT	DISULFID	485	501	BY SIMILARITY.
FT	DISULFID	500	511	BY SIMILARITY.
FT	DISULFID	538	583	BY SIMILARITY.
FT	DISULFID	582	591	BY SIMILARITY.
FT	CONFLICT	27	27	H -> D (IN REF. 6).
FT	CONFLICT	33	33	H -> D (IN REF. 6).
FT	CONFLICT	41	41	Q -> I (IN REF. 6).
SQ	SEQUENCE	608 AA;	68692 MW;	292F7C7ED3A61B4 CRC64;

Query Match 76.6%; Score 1584; DB 1; Length 608;
Best Local Similarity 71.8%; Pred. No. 2.8e-110; Indels 0; Gaps 0;
Matches 278; Conservative 56; Mismatches 53;

QY	1	DAHKEVAHRPKDGEENFKALVLIAPAYLQCPFFEDHVKLVNEVTEFAKTCVADESAAE 60
Db	25	BAHKSEIAHRYNDLGEQHFGLVLIATFQYLOKSCYDEHAKLVQEVTDFAKTCVADESAA 84
QY	61	NCDKSLHLLFGDKLCTVATLRETYGEVADCCAKOEPERNECFLOHKDNNPLALVPEV 120
Db	85	NCDKSLHLLFGDKLCAIENLRENYGELADCTKQEPERNECFLOHKDNNPLPPEPDEA 144
QY	121	DVMCTAFHNDRETEFLKCYLAEIARRHPYFYAPPELLFAKRYKAAATECCQAADKAACLLP 180
Db	145	EMACTSFKENPTTFMGHLYHEVARHPYFYAPPELLYVAEQNYEITLTCCAEADKESCLTP 204
QY	181	KLDELDEGKASSAKORLKCSLOKFGEPRAKAVARLSORFPKAFVSKLVTLTK 240
Db	205	KLDGVKEALVSSVRQRMKSSMOKGERAPKAVARLSOTFFNADFAEITKLATDLTK 264
QY	241	VTECHGDLLECADRADLAKYICENQDSISSKLECEKFLLEKSHCIAEVENDEMPA 300
Db	265	VNKECHGDLLECADRAELAKYMCENQATISSKLTCCDKPELLKKAHCLSEVEDTMPA 324
QY	301	DLPSLAADFVSKDKYKNAEAKDVFLGMFLYEVARRHPDYVSVLLRLAKTYETITLKC 360
Db	325	DLFAIADDFVEDQEVCKNYAEAKDVFLGTFLYYSRRHPDYVSVLLRLAKTYETITLKC 384
QY	361	CAAADPHECYAKVDFEKPVLVEEPNL 387
Db	385	CAENPPACYGTVLAEFQPLVEEPKQL 411

RESULT 13
ALBU CHICK
ID ALBU CHICK STANDARD; PRT; 615 AA.
AC P19121,
DT 01-NOV-1990 (Rel. 15, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).
GN ALB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_taxid=9031;

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=MGS IDR; TISSUE=Liver;
RX  MEDLINE=98116663; PubMed=9455485;
RA  Yoshida K., Seto-Oshima A., Sinohara H.;
RT  "Sequencing of cDNA encoding serum albumin and its extrahepatic
RL  synthesis in the Mongolian gerbil, Meriones unguiculatus.";
RL  DNA Res 4:351-354(1997).
CC  -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC  binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC  hormones, bilirubin and drugs. Its main function is the regulation
CC  of the colloidal osmotic pressure of blood.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- TISSUE SPECIFICITY: Plasma.
CC  -1- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC  -1- SIMILARITY: Contains 3 albumin domains.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AB006197; BAA21765.1; -.
CC  DR  PIR; JCS838; JCS838.
CC  DR  HSP; P02768; 1E7B.
CC  DR  InterPro; IPR000264; Serum_albumin.
CC  DR  Pfam; PF00273; Transport_Prot; 3.
CC  DR  PRINTS; PD00802; SERUMALBUMIN.
CC  DR  PRODOM; PD002486; Serum_albumin; 1.
CC  DR  SMART; SM00103; ALBUMIN; 3.
CC  DR  PROSITE; PS00212; ALBUMIN; 3.
KW  Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT  SIGNAL 1 18 BY SIMILARITY.
FT  PROPEP 19 24 BY SIMILARITY.
FT  CHAIN 25 609 SERUM ALBUMIN.
FT  FT  DOMAIN 25 206 ALBUMIN 1.
FT  FT  DOMAIN 213 398 ALBUMIN 2.
FT  FT  DOMAIN 405 596 ALBUMIN 3.
FT  FT  METAL 28 28 COPPER.
FT  FT  METAL 78 87 BY SIMILARITY.
FT  FT  DISULFID 100 116 BY SIMILARITY.
FT  FT  DISULFID 115 126 BY SIMILARITY.
FT  FT  DISULFID 149 194 BY SIMILARITY.
FT  FT  DISULFID 193 202 BY SIMILARITY.
FT  FT  DISULFID 225 271 BY SIMILARITY.
FT  FT  DISULFID 270 278 BY SIMILARITY.
FT  FT  DISULFID 290 304 BY SIMILARITY.
FT  FT  DISULFID 303 314 BY SIMILARITY.
FT  FT  DISULFID 341 386 BY SIMILARITY.
FT  FT  DISULFID 385 394 BY SIMILARITY.
FT  FT  DISULFID 417 463 BY SIMILARITY.
FT  FT  DISULFID 462 473 BY SIMILARITY.
FT  FT  DISULFID 486 502 BY SIMILARITY.
FT  FT  DISULFID 501 512 BY SIMILARITY.
FT  FT  DISULFID 539 584 BY SIMILARITY.
FT  FT  DISULFID 583 592 BY SIMILARITY.
FT  SEQUENCE 609 AA; 68940 MW; 9CA5F97F67E7F1A48 CRC64;
Query Match 76.6%; Score 1585; DB 1; Length 609;
Best Local Similarity 73.6%; Pred. No. 2.4e-110;
Matches 284; Conservative 40; Mismatches 62; Indels 0; Gaps 0;
QY 2 AKHSEVAHFKDLGEENFKALVLIAPAOYLQCPEDHVKLVNTEPAKTCVADESARN 61
DB 27 AKHSELAHFKDLGEGYFKGLVLYTFSQYLQCSVEEHVKLVREYVDFASNCADKESARN 86
QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFQHKDDNPNLRLVRPEVD 121
DB 87 CDKSLHTLFGDKLCSLPNGEYKAEMADCCAKQEPERNECFQHKDDNPNLPPFKRAEPD 146

```

```

QY 122 VWCATFHDNBEETFLKYLVEIARRHPYFYAPPELLFFANRYKAAFTCCQADKAAALLPK 181
DB 147 AMCTAFQENAEAFMGHYLHEVARHPHYFYGFPELLYLADKYTAVLTTECCADDKACLTPK 206
QY 182 LDELRDGGKASSAKORUKCASLOKFGERAFAKANAVARLSORFPKABEAEVSKLVTDLTKV 241
DB 207 LDALKKALYSANRQRLKCSMKFGERAFAKANAVARMSQTFPNADPAETIKLATDITKV 266
QY 242 HTECHGDLLECADRADLAKYICENODSISKKKECEKPLEKSHCIAEVENDEMPAD 301
DB 267 TQECHGDLLECADRAELAKYMCENQASISSKLQACCKEMQLQSKSCLAEVEHDDMPAD 326
QY 302 LPSLAADPVSQVKYNAEAKOVFLQMFLYEVARHPDYVSVLLLELAKTYETLEKCC 361
DB 327 LPALTADFVEDKVKYNAEAKOVFLCTFLYFSRRHPHYSVSLLLELAKTYETLEKCC 386
QY 362 AAADPHECYAKVDFEKPPLVEEPQNL 387
DB 387 AEADPHACYGHVDFEKPPLVEEPQNL 412
RESULT 12
ALBU_MOUSE
ID ALBU_MOUSE STANDARD; PRT; 608 AA.
AC P07724; Q51802;
DT 01-APR-1988 (Rel. 07, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB OR ALBI OR ALB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
EX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baran G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang X.H., Weitz C., Whittaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
EX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

```

QY 1 DAHSEVAHRPKDLGEENFKALVLIAPAOYLQOCPPEDHVKLVNVEVTEFAKTCVADESAAE 60
 DB 25 EAHKSEIAHRPKDLGEENFKALVLIAPAOYLQOCPPEDHVKLVNVEVTEFAKTCVADESAAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGMAADCCAKQEPERNECFLOHKDDNPNLPRIVREPV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGMAADCCAKQEPERNECFLOHKDDNPNLPRIVREPV 144
 QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 180
 DB 145 EAMCTSFQENPTSLFGLHYLHEVARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 204
 QY 181 KLDELDEGKASSAKQRLKASQKGERAFKAWAVARLSQRPFAEAEVSKLVTDLTK 240
 DB 205 KLDAVEKALVAARQKMCSSQRPGERAFKAWAVARLSQRPFAEAEVSKLVTDLTK 264
 QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKKECEKPLEKSHCIAEVENDEMPA 300
 DB 265 INKECHGDLLECADRADLAKYICENQDSISSKKECEKPLEKSHCIAEVENDEMPA 324
 QY 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYVARRHPDYSVLLRLRLAKTYETTLK 360
 DB 325 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYVARRHPDYSVLLRLRLAKTYETTLK 384
 QY 361 CAADPHCEYAKVDFEFPKLVPEEPQNL 387
 DB 385 CAEGDPACVGTVALEFPQLVPEEPQNL 411

RESULT 10
 ALBU RABIT
 ID ALBU RABIT STANDARD; PRT; 608 AA.
 AC P49065;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.

ALB
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=New Zealand white; TISSUE=Liver;
 RA Sheffield W.P., Syed S., Schuyler P.D.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U18344; AAB58347.1; -;
 DR HSPF; P02768; 1E7B.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.

FT SIGNAL 1 18 BY SIMILARITY.
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIN 25 608 SERUM ALBUMIN.
 FT 25 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.
 FT 404 595 ALBUMIN 3.
 FT METAL 27 27 COPPER.
 FT 77 86 BY SIMILARITY.
 FT DISULFID 99 115 BY SIMILARITY.
 FT 114 125 BY SIMILARITY.
 FT DISULFID 148 193 BY SIMILARITY.
 FT 192 201 BY SIMILARITY.
 FT DISULFID 224 270 BY SIMILARITY.
 FT 269 277 BY SIMILARITY.
 FT DISULFID 289 303 BY SIMILARITY.
 FT 302 313 BY SIMILARITY.
 FT DISULFID 340 385 BY SIMILARITY.
 FT 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.
 FT 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT 582 591 BY SIMILARITY.
 SQ SEQUENCE 608 AA; 68914 MW; CF5E92647AAFE9A2 CRC64;

Query Match 76.6%; Score 1585; DB 1; Length 608;
 Best Local Similarity 71.1%; Pred. No. 2.4e-110;
 Matches 275; Conservative 57; Mismatches 55; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDLGEENFKALVLIAPAOYLQOCPPEDHVKLVNVEVTEFAKTCVADESAAE 60
 DB 25 EAHKSEIAHRPKDLGEENFKALVLIAPAOYLQOCPPEDHVKLVNVEVTEFAKTCVADESAAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGMAADCCAKQEPERNECFLOHKDDNPNLPRIVREPV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGMAADCCAKQEPERNECFLOHKDDNPNLPRIVREPV 144
 QY 121 DVNCTAFHNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 180
 DB 145 DVLCAPFHDEDEKAFGHVLYVARRHPYFAPPELLFFAKRYKAAFTCCOAAADKAAACLLP 204
 QY 181 KLDELDEGKASSAKQRLKASQKGERAFKAWAVARLSQRPFAEAEVSKLVTDLTK 240
 DB 205 KLDALEGGKSLISAQAQERLCASIQKFGDRAYKAWAVARLSQRPFAEAEVSKLVTDLTK 264
 QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKKECEKPLEKSHCIAEVENDEMPA 300
 DB 265 VHTCECHGDLLECADRADLAKYICENQDSISSKKECEKPLEKSHCIAEVENDEMPA 324
 QY 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYVARRHPDYSVLLRLRLAKTYETTLK 360
 DB 325 GLPVAEAEFVEDKQVCKNYEAKOLFLGKFLYVARRHPDYSVLLRLRLAKTYETTLK 384
 QY 361 CAADPHCEYAKVDFEFPKLVPEEPQNL 387
 DB 385 CATDPPHACYAKVDFEFPKLVPEEPQNL 411

RESULT 11
 ALBU MERUN
 ID ALBU MERUN STANDARD; PRT; 609 AA.
 AC C35090;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Meriones.
 OX NCBI_TaxID=10047;

QY 1 DAHKEVAHFKDILGEENFKALVLIAPAOYLQCPEDHVKLVNVEYFAKTCVADSAE 60
 DB 25 DTHKEIAHFNLDGEENFQGLVLIAPSOYLQCPEDHVKLVNVEYFAKTCVADSAE 84
 QY 61 NCDKSLHTLFGDKLCTVATREYVGENMADCCAKQEPERNECFQHKDDNPRLVRPEV 120
 DB 85 GCDKSLHTLFGDELCKVATLREYVGMADCCCEQEPERNECFLNHKDDSPDLPLK- KPFP 143
 QY 121 DVNCTAFHNEETFLKKLYIEIARRHPFYAPELLFFAKRYKAAFTCCQADKAACLLP 180
 DB 144 DTLCAEFKADKKFPMGKLYLEVARRHPFYAPELLFYANKYNGVFOCCQADKGAACLLP 203
 QY 181 KLDELREGKASAKORLKCASLQKGERAFKAWAVARLSORPKAEFAEVSKLVTDLTK 240
 DB 204 KIDAMREKVLASSARQLRCASLQKGERALKAWVARLSQKPKADFTDVTKIVTDLTK 263
 QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCECKPLLEKSHCIAEVENDEMPA 300
 DB 264 VHKECCHGDLLECADRADLAKYICDHQDLSLSEKLEKCECKPLLEKSHCIAEVDKDAVPE 323
 QY 301 DLPSLAADFVESKDVCKNVAEADVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360
 DB 324 NLPLTADPADFEDKEVCKNQYQKADVFLGSLFYSSRRHPYAVSVLRLAKYETLEDC 383
 QY 361 CAAADPHCYAKVDFEKLVEFPQNL 387
 DB 384 CAKEDPHACYATVDFKLHLVDFPQNL 410

RESULT 9

ALBU RAT STANDARD; PRT; 608 AA.
 AC P02770; F11382;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
 GN ALB.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8123722; PubMed=7017712;
 RA Sargent T.D., Yang M., Bonner J.;
 RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
 RN [2]
 RP SEQUENCE OF 1-38, AND PROCESSING.
 RX MEDLINE=77249657; PubMed=893447;
 RA Straus A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
 RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
 RT piece. Analysis of the direct translation product of albumin
 RT messenger RNA.";
 RL J. Biol. Chem. 252:6846-6855(1977).
 RN [3]
 RP SEQUENCE OF 25-222.
 RX MEDLINE=78109429; PubMed=564345;
 RA Isemura S., Ikenaka T.;
 RT "Amino acid sequences of fragments I and II obtained by cyanogen
 RT bromide cleavage of rat serum albumin.";
 RL J. Biochem. 83:35-48(1978).
 RN [4]
 RP SEQUENCE OF 223-288 AND 572-608.
 RX MEDLINE=78260153; PubMed=956149;
 RA Isemura S., Ikenaka T.;
 RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
 RT the amino acid sequences of four fragments.";
 RL J. Biochem. 79:1183-1196(1976).
 RN [5]
 RP SEQUENCE OF 166-174.
 RC TISSUE=Plasma;

RX MEDLINE=87194805; PubMed=2437111;
 RA Caraway R.E., Mitra S.P., Cochran D.E.;
 RT "Structure of a biologically active neurotensin-related peptide
 RT obtained from pepsin-treated albumin(s).";
 RL J. Biol. Chem. 262:5968-5973(1987).
 RN [6]
 RP COPPER-BINDING.
 RX MEDLINE=79001617; PubMed=80265;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";
 RL Cancer Res. 38:3483-3486(1978).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- FUNCTION: NRP regulates fat digestion, lipid absorption, and
 CC blood flow (potential).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch)
 CC -----
 CC EMBL: V01222; CAA24532.1; -.
 DR PIR: A93872; ABRYS.
 DR HSP: P02768; IE7B.
 DR InterPro: IPR00264; Serum albumin.
 DR Pfam: PF00273; transport prot; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum albumin; 1.
 DR SMART: SM00103; ALBUMIN_3.
 DR PROSITE: PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT PEPTIDE 156 174
 FT DOMAIN 212 397
 FT DOMAIN 404 595
 FT METAL 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 148 193
 FT DISULFID 192 201
 FT DISULFID 224 270
 FT DISULFID 269 277
 FT DISULFID 289 303
 FT DISULFID 302 313
 FT DISULFID 340 385
 FT DISULFID 384 393
 FT DISULFID 416 462
 FT DISULFID 461 472
 FT DISULFID 485 501
 FT DISULFID 500 511
 FT DISULFID 538 583
 FT DISULFID 582 591
 FT VARIANT 282 262
 FT CONFLICT 174 174
 FT SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;
 Y -> L (IN REF. 5).
 Y 77.9%; Score 1610; DB 1; Length 608;
 Best Local Similarity 72.1%; Pred. No. 3.4e-112;
 Matches 279; Conservative 59; Mismatches 49; Indels 0; Gaps 0;

Query Match

77.9%; Score 1610; DB 1; Length 608;

Best Local Similarity 72.1%; Pred. No. 3.4e-112;

Matches 279; Conservative 59; Mismatches 49; Indels 0; Gaps 0;

```
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
KW Polymorphism. 1 18
FT SIGNAL 19 24 SERUM ALBUMIN.
FT PROPEP 25 607
FT CHAIN 25 204 ALBUMIN 1.
FT DOMAIN 25 396
FT DOMAIN 211 396
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
FT DISULFID 590 607
FT VARIANT 214 214
FT CONFLICT 302 302
FT CONFLICT 304 304
FT CONFLICT 324 324
FT CONFLICT 324 324
FT CONFLICT 394 395
FT CONFLICT 437 437
FT CONFLICT 437 437
FT CONFLICT 493 494
SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;

Query Match 78.5%; Score 1622.5; DB 1; Length 607;
Best Local Similarity 75.2%; Pred. No. 4e-113;
Matches 291; Conservative 46; Mismatches 49; Indels 1; Gaps 1;

QY 1 DAHSEVAHFKDLGEBNFKALVLIAPQVLOQCPEDHVKLVNTEFAKTCVADESAE 60
DB 25 DTHKSEIAHFKDLGEBNFKALVLIAPQVLOQCPEDHVKLVNTEFAKTCVADES 84
QY 61 NCKSLHTLFGDKLCTVATRETYGEMADCCAKQEPERNECFQHKDDNPNRLVRP 120
DB 85 GCKSLHTLFGDELCKVASLRETYGEMADCCQEPERNECFSLHKKDDSPDLK 143
QY 121 DVMCTAFHDNEETLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKA 180
DB 144 NTLCDPEKADKKFKWGLYKLYEIARRHPYFYAPPELLFYANKYNGVFOCCQAE 203
QY 181 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSORPPKAEFVSKLV 240
DB 204 KIETMRKVLASSARQLRCSLQKFGERALKAWAVARLSQKPKAEFVETKLV 263
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKLECKECPKLLKSHCIAVEN 300
DB 264 VHKCECHGDLLECADRADLAKYICDNQDTISSKLECKECPKLLKSHCIAVE 323
QY 301 DLPSLAADFVESKDVCKNYAEKDVFLGMFLYEVAREHPDYVSVLLRLAKTY 360
DB 324 NLPLTADPFDKDVCKNYAEKDAFLGSLFYYSRRHPDYVSVLLRLAKTY 383
QY 361 CAADPHECYAKVDFEKPVLVEEPQNL 387
DB 384 CAKDDPRACYSTVDFKLVDFEPQNL 410

RESULT 8
ALBU_SHEEP STANDARD; PR7; 607 AA.
ID ALBU_SHEEP
AC P14639;
DT 01-APR-1990 (Rel. 14, Created)
```

01-APR-1990 (Rel. 14, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Serum albumin precursor.
ALB.
Ovis aries (Sheep).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprinae; Ovis.
NCBI_TaxID=9940;
[1]
SEQUENCE FROM N.A.
TISSUE=Liver;
MEDLINE=9009888; PubMed=2602160;
Brown W.M., Dziegielewska K.M., Foreman R.C., Saunders N.R.;
"Nucleotide and deduced amino acid sequence of sheep serum albumin";
Nucleic Acids Res. 17:10495-10495(1989).
-!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
-!- SUBCELLULAR LOCATION: Secreted.
-!- TISSUE SPECIFICITY: Plasma.
-!- SIMILARITY: Belongs to the ALB/APP/VDB family.
-!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).

EMBL; X17055; CAA34903.1; -.
PIR; S06936; ABSHS.
DR HSP; P02768; 1E7B.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PS00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 607
FT DOMAIN 25 204
FT DOMAIN 211 396
FT DOMAIN 403 594
FT METAL 27 27
FT METAL 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 147 192
FT DISULFID 191 200
FT DISULFID 223 269
FT DISULFID 268 276
FT DISULFID 288 302
FT DISULFID 301 312
FT DISULFID 339 384
FT DISULFID 383 392
FT DISULFID 415 461
FT DISULFID 460 471
FT DISULFID 484 500
FT DISULFID 499 510
FT DISULFID 537 582
FT DISULFID 581 590
SQ SEQUENCE 607 AA; 69188 MW; 84979A87P8B86596 CRC64;
Query Match 78.0%; Score 1613.5; DB 1; Length 607;
Best Local Similarity 74.2%; Pred. No. 1.8e-112;
Matches 287; Conservative 48; Mismatches 51; Indels 1; Gaps 1;


```

ID ALBU PELCA STANDARD; PRT; 608 AA.
AC P49064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Fel d 2).
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96194824; PubMed=8647469;
RA Hilger C., Grigioni F., Kohnen M., Hentges F.;
RT "Sequence of the gene encoding cat (Felis domesticus) serum albumin.";
RL Gene 169:295-296(1996).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- ALLERGEN: Causes an allergic reaction in human.
CC -1- SIMILARITY: Belongs to the ALB/AFB/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X84842; CAA59279.1; -.
DR PIR; JC4660; S57632.
DR HSP; P02768; 1E7B.
DR InterPro; IPR00264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PD00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT SIGNAL 1 18
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 608 SERUM ALBUMIN.
FT DOMAIN 25 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 27 27 COPPER.
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 340 383 BY SIMILARITY.
FT DISULFID 384 395 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
SQ SEQUENCE 608 AA; 68659 MW; 07E629CAC5F6085F CRC64;
Query Match 83.5%; Score 1727; DB 1; Length 608;
Best Local Similarity 80.6%; Pred. No. 7e-121;

```

```

Matches 312; Conservative 34; Mismatches 41; Indels 0; Gaps 0;
QY 1 DAHSEVAHREPKDLGEENFKALVLIAPAOVLOQCPEFHVKLVNEVTEPAKTCVADSEAE 60
DB 25 EAHQSEIAHREPNLGEHFRGLVLAFAVSQYLOQCPEFHVKLVNEVTEPAKTCVADSEAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
DB 85 NCEKSLHELLGDKLCTVASLRDKYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 144
QY 121 DVMTAFHDNEBETPLKYLIEIARRHYPFYAPDELLFFAKRYKAAFTCCCAADKAAACLLP 180
DB 145 DAMCTAFHENEQRFLKYLIEIARRHYPFYAPDELLFFAKRYKAAFTCCCAADKAAACLLP 204
QY 181 KLDELREGKASSAKQRLKCSAQKQGERAFKAWAVARLSORPPKAFVAVSLVLTDLTK 240
DB 205 KVDALREKYLASSAKERLKCSAQKQGERAFKAWAVARLSORPPKAFVAVSLVLTDLTK 264
QY 241 VHTCCGHDLLLECAADRADLAKYICENQDSISSKLKECCXPLEKHCHIAEVENDEMPA 300
DB 265 IHKCCGHDLLLECAADRADLAKYICENQDSISTKLKECCGKPVLEKSHCHISEVERDELPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFIEYARRHPDYSVVLRLRLAKTVETTLK 360
DB 325 DLPLAVDFVEDKVCNKYQKADVFLGTFLYEYRRHPEYSVLLRLAKVEZATLEKC 384
QY 361 CAADPHCEYAKVDFEFPKPLVEEPQL 387
DB 385 CATDDPPACYAHVDFEFPKPLVEEPHNL 411
RESULT 4
ALBU CANFA STANDARD; PRT; 608 AA.
ID ALBU CANFA STANDARD; PRT; 608 AA.
AC P49822; O77705; Q9TSZ4;
DT 01-OCT-1996 (Rel. 34, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Allergen Can f 3).
GN ALB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Liver;
RA Hilger C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20148667; PubMed=10669848;
RA Pandjaitan B., Swoboda I., Brandesky-Fichler F., Rumpold H.,
RA Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant dog
RT albumin, a cross-reactive animal allergen.";
RL J. Allergy Clin. Immunol. 105:279-285(2000).
RN [3]
RP SEQUENCE OF 25-48.
RX MEDLINE=75011422; PubMed=4414013;
RA Dixon J.W., Sarkar B.;
RT "Isolation, amino acid sequence and copper(II)-binding properties of
RT peptide (1-24) of dog serum albumin.";
RL J. Biol. Chem. 249:5872-5877(1974).
RN [4]
RP SEQUENCE OF 25-38.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504912;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).

```

Query Match 100.0%; Score 2068; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 3.5e-146;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHFRFDLGEENFKALVLIAPAOYLQCCPEEDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSVAHFRFDLGEENFKALVLIAPAOYLQCCPEEDHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 144
QY 121 DVMTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAACLLP 180
DB 145 DVMTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAACLLP 204
QY 181 KLDELDEGRKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVDTLTK 240
DB 205 KLDELDEGRKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVDTLTK 264
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECCPKLLEKSHCIAEVENDEMPA 300
DB 265 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECCPKLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADPVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVLLRLAKTYETTLTKC 360
DB 325 DLPSLAADPVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVLLRLAKTYETTLTKC 384
QY 361 CAADPHECYAKVDFBPKPLVEEPQNL 387
DB 385 CAADPHECYAKVDFBPKPLVEEPQNL 411

RESULT 2
ALBU_MACMU STANDARD; PRT; 600 AA.
ID ALBU_MACMU
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93211971; PubMed=8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwulet J., Putnam F.W.;
RT "cDNA and protein sequence of polymorphic macaque albumins that differ
in bilirubin binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: Belongs to the ALB/AFB/VDB family.
CC -1- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; M90463; AAA36906.1; -.
DR FIR; A47391; A47391.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot. 3.
DR PRINTS; PS00802; SERUMALBUMIN.
DR PRODOM; PD002486; Serum albumin; 1.
DR SMART; SMO0103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT NON_TER 1 1
FT SIGNAL <1 10 BY SIMILARITY.
FT PROPEP 11 16 BY SIMILARITY.
FT CHAIN 17 600 SERUM ALBUMIN.
FT DOMAIN 17 197 ALBUMIN 1.
FT DOMAIN 204 389 ALBUMIN 2.
FT DOMAIN 396 587 ALBUMIN 3.
FT METAL 19 19 COPPER (BY SIMILARITY).
FT BINDING 256 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; E45C871A670E740B CRC64;

Query Match 96.1%; Score 1988; DB 1; Length 600;
Best Local Similarity 94.6%; Pred. No. 2.9e-140;
Matches 366; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

QY 1 DAHKSVAHFRFDLGEENFKALVLIAPAOYLQCCPEEDHVKLVNEVTEFAKTCVADESAAE 60
DB 17 DTHKSEVAHFRFDLGEENFKGLVLAFAVSQYLQCCPEEHVKLVNEVTEFAKTCVADESAAE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 136
QY 121 DVMTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAACLLP 180
DB 137 DVMTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAACLLP 196
QY 181 KLDELDEGRKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVDTLTK 240
DB 197 KLDELDEGRKASSAKORLKCASLQKFGERAFAKAWAVARLSQRPFAEFAEVSCLVDTLTK 256
QY 241 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECCPKLLEKSHCIAEVENDEMPA 300
DB 257 VHTTECHGDLLECCADRADLAKYICENQDSISSKLECCPKLLEKSHCIAEVENDEMPA 316
QY 301 DLPSLAADPVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVLLRLAKTYETTLTKC 360
DB 317 DLPSLAADPVESKDVCKNVAEAKDVLGMFLVEYARRHPDYSVLLRLAKTYETTLTKC 376
QY 361 CAADPHECYAKVDFBPKPLVEEPQNL 387
DB 377 CAADPHECYAKVDFBPKPLVEEPQNL 403

RESULT 3
ALBU_FELCA

RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Tothiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Hilton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield A.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RN SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RN SEQUENCE OF 25-609.
RA Brown J.R., Shockley P., Behrens P.O.;
RL (In) Bing D.H. (eds.);
RL The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RN SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11]
RN SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tanaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RN SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RN SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
1994.";
RL Electrophoresis 15:1459-1465(1994).
RN [14]
RN DISULFIDE BONDS.
RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15]
RN BIIRUBIN-BINDING SITE.
RA Jacobsen C.;
RX MEDLINE=78186630; PubMed=656055;
RT "Lysine residue 240 of human serum albumin is involved in high-
affinity binding of bilirubin.";
RL Biochem. J. 171:453-459(1978).

RN [16]
RN VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herber P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RN VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=88068523; PubMed=3479777;
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satch C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
Amerindian and Japanese populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18]
RN VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huss K., Ishioke N., Satch C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese alloalbumins.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19]
RN VARIANTS MANAUS; OSAKA; NAGOVA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20]
RN DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
human serum albumin whose precursor has an aberrant signal peptidase
cleavage site.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21]
RN VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
Italy.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22]
RN VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
carboxyl-terminal variants of human serum albumin.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23]
RN VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RX MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Anaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN [24]
RN VARIANT CASEBROOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
albumin: albumin Casebrook (494 Asp-->Asn).";
RL Biochim. Biophys. Acta 1097:49-54(1991).
RN [25]
RN VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two alloalbumins with identical electrophoretic mobility are produced

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 28.2299 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18_COPY_1_387
Perfect score: 2068
Sequence: 1 DAHSEVAHRFKDLEENFX.....ECYAKVDFEKLVEEPQNL 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2068	100.0	609	1	ALBU HUMAN
2	1998	96.1	600	1	ALBU MACMU
3	1727	83.5	608	1	ALBU FELCA
4	1690	81.7	608	1	ALBU CANFA
5	1660	80.3	607	1	ALBU HORSE
6	1623.5	78.5	605	1	ALBU PIG
7	1622.5	78.5	607	1	ALBU BOVIN
8	1613.5	78.0	607	1	ALBU SHEEP
9	1610	77.9	608	1	ALBU RAT
10	1585	76.6	608	1	ALBU RABIT
11	1585	76.6	609	1	ALBU MERUN
12	1584	76.6	608	1	ALBU MOUSE
13	1038.5	50.2	615	1	ALBU CHICK
14	766	37.0	607	1	ALB2 XENLA
15	726.5	35.1	606	1	ALB1 XENLA
16	724.5	35.0	609	1	FETA PANTR
17	721.5	34.9	609	1	FETA GORGO
18	720.5	34.8	609	1	FETA HUMAN
19	691	33.4	599	1	AFAM HUMAN
20	675	32.6	609	1	FETA HORSE
21	671	32.4	611	1	AFAM MOUSE
22	656	31.7	608	1	AFAM RAT
23	615	29.8	605	1	FETA MOUSE
24	603	29.2	611	1	FETA RAT
25	468	22.6	608	1	ALB1_SALSA
26	465	22.5	608	1	ALB2_SALSA
27	357	17.3	474	1	VTDB HUMAN
28	353	17.1	382	1	ALBU RANCA
29	348	16.8	476	1	VTDB RAT
30	345	16.7	476	1	VTDB RABIT
31	341	16.5	476	1	VTDB MOUSE
32	311	15.0	1423	1	ALBU_PETWA
33	133	6.4	1605	1	RRB1_MOUSE

34 122.5 5.9 344 1 MST1_DROHY
35 111 5.4 40 1 ALB1_TRASC
36 108 5.2 1410 1 RRB1_HUMAN
37 107.5 5.2 1228 1 YHV5_YEAST
38 106 5.1 650 1 BM86_BOOMI
39 102 4.9 1391 1 MST2_DROHY
40 101.5 4.9 3038 1 TRIO_HUMAN
41 100 4.8 275 1 MST3_DROHY
42 99.5 4.8 1962 1 MYSA_DROME
43 99 4.8 1285 1 SLI7_ENTHI
44 99 4.8 2867 1 RRB2_PLAIV
45 99 4.8 8797 1 SNE1_HUMAN

Q08695 drosophila
P81188 trachemys s
Q92629 homo sapien
P38851 saccharomyc
F20736 boophilus m
Q08696 drosophila
O75962 homo sapien
O01395 drosophila
P05661 drosophila
P23502 entamoeba h
Q00799 plasmodium
Q8n191 homo sapien

ALIGNMENTS

RESULT 1
ALBU HUMAN
ID ALBU HUMAN STANDARD; PRT; 609 AA.
AC P02768; Q95574; Q13140; Q9P157; Q9P117; Q9UHS3; Q9OUZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Rufner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczak A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within q11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX MEDLINE=82081882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wion K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczak A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).
RC TISSUE=Fetal liver;
RA Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,
RA Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;
RT "Functional prediction of the coding sequences of 121 new genes
RT deduced by analysis of cDNA clones from human fetal liver.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
RX Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.

Search completed: April 19, 2004, 12:02:20
Job time : 48.169 secs

```

Db      265 FIKDCGDNFECWTELEISEHTCOHKDELSTKLEKCCNPLLERTYCIVTLENDVPA 324
Qy      301 DLPSLAADFVESKDVKKNYAEAKDVFGMLFYEVARRHPDVSVLLRLAKTYETLTKC 360
Db      325 ELGKPIETFTDPHVCKEYAKNESFLERISPMOSQETPELSQFLLQSAKEYESLLNKC 384
Qy      361 CAAADPHECYAKVDFEF 377
Db      385 CFSNPFECYKGDNR 401

RESULT 15
JC4258
alpha-fetoprotein precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
C:Accession: JC4258
R:Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczky, A.
Gene 162, 213-220, 1995
A:Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to
A:Reference number: JC4258; MUID:96032345; PMID:7557431
A:Accession: JC4258
A:Molecule type: DNA
A:Residues: 1-609 <NIS>
A:Cross-references: GB:J21916; NID:G841311; PIDN:AA91641.1; PID:G841312
C:Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
c similar properties and structure.
C:Genetics:
A:Gene: afp
A:Map position: 3p
A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 55/2
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: Glycoprotein
E:1-19/Domain: signal sequence #status predicted <SIG>
F:20-609/Product: alpha-fetoprotein #status predicted <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:42,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.0%; Score 724.5; DB 2; Length 609;
Best Local Similarity 36.3%; Pred. No. 2,3e-46;
Matches 142; Conservative 75; Mismatches 167; Indels 7; Gaps 3;

Qy      3 HKSE-----VAHRFKDIGEENFKALVLIAPAYLIQQCFEDHVKLNVNVTFAKTCVADE 57
Db      22 HRNEVYGASILDYSQCYAEINLITLATITFAQFVQEATYKEVSKVMDALTAIEKPTGDE 81
Qy      58 SAENCDSLHTLFDGDKLCTVATRETVMGADCCAKOPERNECFLOHKDDNP-NLPLVL 116
Db      82 QSACGLENLQPAFLFELCREKEILEUKYGH-SDCSQSQSEGRNCFLAHKETPASIPFPQ 140
Qy      117 RPEVDVMCTAPHDNEETFLKYLVEIARRHPYFVAPELLFFAKRYKAAFTCCQAADKAA 176
Db      141 VPEVPTSCAEYEDRETFMNKFYIEIARRHPFLYAPTILLWAARYDKIIPSCCKAENAVE 200
Qy      177 CLLPLKDLDRDGKASAKORLKCASLQKFGERAFKAWAVARLSQRPPKAEFVSKLVT 236
Db      201 CFQTKAATVTKELRESSLLNQHACAVMKNQFTRTFQALTIVTKLSQFKFWNFTETQLVL 260
Qy      237 DLTKVHTECCGHDLLLECADDRAADIAKYICENQDSISSKLKCECKEPLLEKSHCIAEVND 296
Db      261 DVAHVHEHCRCRGDVLDCQDGEKMSYTCQQDTSLSNKITCECKLTTLERGQCIIHAEND 320
Qy      297 EMPADLPSLAADFVESKDVKKNYAEAKDVFGMLFYEVARRHPDVSVLLRLAKTYET 356
Db      321 EKPGCLSPNLNRFICDRDFNQFSSGKKNIFLASFVHEYSRHPQLAVSVILVAKYQEL 380
Qy      357 LEKCAAADPHECYAKVDFEKPFLVBEPNL 387
Db      381 LEKQFTENPLECDKGEEELQYIQESQAL 411

```


A;Cross-references: EMBL:X13060; NID:g52939; PIDN:CAAJ1458.1; PID:g8999334
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F;123-296/Domain: serum albumin repeat homology (<S2>
F;315-453/Domain: serum albumin repeat homology (fragment) <SA3>

Query Match 61.1%; Score 1264; DB 2; Length 453;
Best Local Similarity 70.6%; Pred. No. 1.1e-86;
Matches 221; Conservative 45; Mismatches 47; Indels 0; Gaps 0;

Qy 75 CTVATLTETGEMADCCAKPEPENECLFHQKDNPMLPLRVPRVDVMCTAFHDNEETF 134
Db :
1 CAPNLRENTYGEALDCCTKQEPENECFLQHNDNPSPUPPERPEAEAWCTSFKENPTIF 60
Qy 135 LKKYLVEIARRHPFYPAPELLFFAKRYKAAPTECCQAADKAACLLPKLDLRDEGKASSA 194
Db :
61 MGHYLHEVAARHPFYPAPELLYYAEQNYEILTQCCEADKESCLTPKLGDGVKEKALVSSV 120
Qy 195 KORLKCSLOKFGRAPKANAVAELSORFPKAEFAEVSKLVTDITKYHTECHGLLECA 254
Db :
121 QRWKCSSMQKFGRAKANAVAEVSOTFPNADFITKLATDITKYNKECHGDLLECA 180
Qy 255 DDRADLAKYCENODSISSLKECEKPLEKSHCIAEVENDEMADIPSLAADPFVESKD 314
Db :
181 DDRAELAKWCENOATISSLKTCCDXPLLKKAHCLSEVEDHTWPADLPAAADPFVDEQE 240
Qy 315 VCKRYAAKDVLFQMFLEYEARHPDYSVULLLLAKTYETTLEKCMAAADPHCYAKVF 374
Db :
241 VCKRYAAKADVFLGTFLYEYSRRHPDYSVULLLLAKYEATLEKCCAANPPACYGTVL 300
Qy 375 DEPKPLVSEPONL 387
Db :
301 AEFQPLVEEPKNL 313

RESULT 11
ABCS
Serum albumin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 22-Jun-1999
F;Accession: S15571; A05078; A13451
R;Cassidy, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A;Reference number: S15571
A;Accession: S15571
A;Molecule type: mRNA
A;Residues: 1-615 <CR>
A;Cross-references: EMBL:X60688; NID:g63747; PIDN:CAA43098.1; PID:g63748
R;Hache, R.J.G.; Wiskocil, R.; Vaea, M.; Roy, R.N.; Lau, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A;Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein
A;Reference number: A05078; MUID:83161037; PMID:6187737
A;Accession: A05078
A;Molecule type: DNA
A;Residues: 1-28 <HC>
A;Cross-references: GB:V00381; NID:g63038; PIDN:CAA23680.1; PID:g63039
R;Rosen, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A;Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A;Reference number: A13451; MUID:78019943; PMID:911327
A;Accession: A13451
A;Molecule type: Protein
A;Residues: 19-23; M, 28-30 <ROS>
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mone (weak bonds with these hormones promote their transfer across the membranes), thy
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-26/Domain: propetide #status predicted <PRO>
F;21-613/Product: serum albumin #status predicted <MAT>
F;32-206/Domain: serum albumin repeat homology <SA1>
F;225-398/Domain: serum albumin repeat homology <SA2>

F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-203/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 78.0%; Score 1613.5; DB 1; Length 607;
Best Local Similarity 74.2%; Pred. No. 1.3e-112; Indels 1; Gaps 1;
Matches 287; Conservative 48; Mismatches 51;
QY 1 DAHSEVAHREKDLGEENFKALVLIAPAFYQQLQCPFFDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DTHKSEIARRNDLGEENFQGLVLIAPFQYQQLQCPFFDHVKLVKELTEFAKTCVADESAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 GCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPNLRLVRPEV 143
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 144 DTLCAEFKADKKWGKLYEVARHPYFYAPPELLYANKYNGVFQSCQAEDKGAACLLP 203
QY 181 KLDELDEGKASSAKQRLKCSAQKFGGERAFKAWARLSQRPFKAEFAEYVKLVTLDTLK 240
DB 204 KIDAMREKVLASSARQRLKCSAQKFGGERALKAWARLSQRPFKADFTDVTIVTLDTLK 263
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKKECCPEFLKSHCHIAEYVNDMPA 300
DB 264 VKECCCHGDLLECADRADLAKYICDHODALSSKKECCDPVLEKSHCHIAEYVDKAVPE 323
QY 301 DLPSLAADFVESKDVCKNYAKADVFLGMFLYIYARRHPDYVSVLLRLAKTYETTLK 360
DB 324 NLPLTADFADBEKCKNYQAKDVFLGSLFYIYARRHPDYVSVLLRLAKTYETTLK 383
QY 361 CAADPHCYAKVDFDEKPLVEEPQNL 387
DB 384 KAKEDPHACYATVFDKLHLVDEPQNL 410

RESULT 8

ABRTS
serum albumin precursor - rat
N:Alternate names: preproalbumin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 76, 243-246, 1981
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; MUID:81223722; PMID:7017712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <SAR>
A:Cross-references: GB:V01222; GB:J00698; NID:955627; PIDN:CAA24532.1; PID:955628
R:Stauss, A.W.; Bennett, C.D.; Donchue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
A:Reference number: A92211; MUID:77249657; PMID:893447
A>Note: cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <STR>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A:Reference number: A91946; MUID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein
A:Residues: 25-222 <IS1>
R:Isemura, S.; Ikenaka, T.

J. Biochem. 79, 1183-1196, 1976
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino
A:Reference number: A91940; MUID:76260153; PMID:956149
A:Accession: A91940
A:Molecule type: protein
A:Residues: 223-288;572-608 <IS2>
A>Note: 262-Leu was also found
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.
A:Reference number: A90758; MUID:79001617; PMID:80265
A:Contents: annotation; copper binding
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid protease
A:Reference number: A45800; MUID:89341406; PMID:2474609
A:Accession: C45800
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 166-173 <CAR>
R:Heard, J.

Mol. Cell. Biol. 7, 2425-2434, 1987
A:Title: Determinants of rat albumin promoter tissue specificity analyzed by an improv
A:Reference number: I57621; MUID:87286876; PMID:3475566
A:Accession: I57621
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5 <RES>
A:Cross-references: GB:M16825; NID:G202828; PIDN:AAA40712.1; PID:G554412
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status experimental <PRO>
F:19-24/Domain: propeptide #status experimental <PRO>
F:25-608/Product: serum albumin #status experimental <MAT>
F:29-202/Domain: serum albumin repeat homology <SA1>
F:221-394/Domain: serum albumin repeat homology <SA2>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status experimental
F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,

Query Match 77.9%; Score 1610; DB 1; Length 608;

Best Local Similarity 72.1%; Pred. No. 2.4e-112;

Matches 279; Conservative 59; Mismatches 49; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGEENFKALVLIAPAFYQQLQCPFFDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 EAHKSEIARRNDLGEENFQGLVLIAPFQYQQLQCPFFDHVKLVKELTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 EAMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSAQKFGGERAFKAWARLSQRPFKAEFAEYVKLVTLDTLK 240
DB 205 KLDVKEKALVAARQMKCSQVQFGERAFKAWARLSQRPFKAEFAEYVKLVTLDTLK 264
QY 241 VHTCECHGDLLECADRADLAKYICENQDSISSKKECCPEFLKSHCHIAEYVNDMPA 300
DB 265 INKECCCHGDLLECADRADLAKYICENQDSISSKKECCPEFLKSHCHIAEYVNDMPA 324
QY 301 DLPSLAADFVESKDVCKNYAKADVFLGMFLYIYARRHPDYVSVLLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAKADVFLGMFLYIYARRHPDYVSVLLRLAKTYETTLK 384
QY 361 CAADPHCYAKVDFDEKPLVEEPQNL 387
DB 385 CAEGDPPACVGTVLAEFQPLVEEPQNL 411

C;Accession: A38885; A36401; A91258; B60808; S10780; D45800; A26693; A90309; A91458; A94
R;Hollowachuk, E.W.; Stoltenberg, J.K.; Reed, R.G.; Peters Jr., T.
submitted to the EMBL Data Library, August 1991
A;Description: Bovine serum albumin: cDNA sequence and expression.
A;Reference number: A38885
A;Accession: A38885
A;Molecule type: mRNA
A;Residues: 1-607 <HOL>
A;Cross-references: EMBL:M73215
R;Hirayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
A;Title: Rapid confirmation and revision of the primary structure of bovine serum albumin
A;Reference number: A36401; MUID:91083649; PMID:2260975
A;Accession: A36401
A;Molecule type: protein
A;Residues: 25-41, 'H', 43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>
R;MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
Eur. J. Biochem. 98, 477-489, 1979
A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A;Reference number: A91258; MUID:80024278; PMID:489109
A;Accession: A91258
A;Molecule type: protein
A;Residues: 1-32 <MAG>
R;Haieh, J.C.; Lin, F.P.; Tam, M.F.
Anal. Biochem. 170, 1-8, 1988
A;Title: Electrophoretic onto glass-fiber filter from an analytical isoelectrofocusing gel
A;Reference number: A60808; MUID:88267456; PMID:3389500
A;Accession: B60808
A;Molecule type: protein
A;Residues: 25-41 <HSI>
R;Strawich, E.; Glimcher, M.J.
Eur. J. Biochem. 191, 47-56, 1990
A;Title: Tooth 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albumin
A;Reference number: S10780; MUID:90336641; PMID:2379503
A;Accession: S10780
A;Molecule type: protein
A;Residues: 25-41, 'H', 43-57, 59-64 <STR>
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1580-1584, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteases
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: D45800
A;Molecule type: protein
A;Residues: 163-172 <CAR>
R;Carraway, R.E.; Mitra, S.P.; Cochran, D.E.
J. Biol. Chem. 262, 5968-5973, 1987
A;Title: Structure of a biologically active neurotensin-related peptide obtained from pe
A;Reference number: A26693; MUID:87194805; PMID:2437111
A;Accession: A26693
A;Molecule type: protein
A;Residues: 165-172, 'L', <CA2>
R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.
Biochem. J. 191, 867-868, 1980
A;Title: Sequence of residues 400-403 of bovine serum albumin.
A;Reference number: A90309; MUID:82023364; PMID:7283978
A;Accession: A90309
A;Molecule type: protein
A;Residues: 402-433 <REE>
R;Brown, J.R.
Fed. Proc. 34, 591, 1975
A;Title: Structure of bovine serum albumin.
A;Reference number: A91458
A;Accession: A91458
A;Molecule type: protein
A;Residues: 25-41, 'H', 43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 196-213, 'T', 215-288, 'E'
R;Brown, J.R.
submitted to the Atlas, April 1975
A;Reference number: A94551
A;Accession: A94551
A;Molecule type: protein
A;Residues: 190-195 <BR2>
R;Brown, J.R.
Fed. Proc. 33, 1389, 1974

A;Reference number: A91457
A;Contents: annotation; disulfide bonds
R;Werlen, R.C.; Offord, R.E.; Rose, K.
Biochem. J. 302, 907-911, 1994
A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC
A;Reference number: S55232; MUID:95031935; PMID:7945219
A;Accession: S55232
A;Status: Preliminary
A;Molecule type: protein
A;Residues: 529-536;569-572 <WER>
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; copper binding; duplication; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-607/Product: serum albumin #status experimental <MPT>
F;29-201/Domain: serum albumin repeat homology <SA1>
F;220-393/Domain: serum albumin repeat homology <SA2>
F;412-591/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,
Query Match 78.3%; Score 1618.5; DB 1; Length 607;
Best Local Similarity 74.9%; Pred No. 5.6e-113;
Matches 290; Conservative 46; Mismatches 50; Indels 1; Gaps 1;
QY 1 DAHSEVAHRFKDLGEEENFKALVLIAPQYLLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DTHKSEIAHRRFKDLGEEQFKGLVLIAPQYLLQCCPFEDHVKLVNELTEFAKTCVADESAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPENECEFLQHKDNPPLRVREPV 120
Db 85 GCKSLHTLFGDECKVASLRETYGDMADCCCKOEPENECEFLSHRDDSPDPKL-KDDP 143
QY 121 DVMCTAFPHDNBEETFLKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAKACLLP 180
Db 144 NTLCDERKADKKFKWGLYIEIARRHPYFYAPPELLFYANKYGVFQDCQAEKDGACLLP 203
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAPKAWAVARLSORFPKAEFAEYVKLVTDLTK 240
Db 204 KIETVREKVLASSARQURLCASIOKFGERALKAVSARLSOKFKPAEFVEVTKLVTDLTK 263
QY 241 VHTCECHGDLLECDRADLAKYICENQDISSSKLKECEKFLLEKSHCIAEVENDEMPA 300
Db 264 VHEKCHGDLLECDRADLAKYICDNQDTISSKLKECCDKFLLEKSHCIAEVEKDAIPE 323
QY 301 DLPSLAADFVSKQVCKYAKQVFLGMFLYEVARHPDYVSVLLRLAKTYETTLK 360
Db 324 NLPLPLTADFEDKQVCKNYQEAQDAFLGSLFYYSRRHPYAVSVLLRLAKYEATLEEC 383
QY 361 CAADPHCEYAKVDPDERKPLVEEPQNL 387
Db 384 CAKDPPHACYSTVFDKLKHLVDFQNL 410
RESULT 7
ABSHS
serum albumin precursor - sheep
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C;Accession: S06936
R;Brown, W.M.; Dziedzielska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A;Reference number: S06936; MUID:90098888; PMID:2602160
A;Accession: S06936
A;Molecule type: mRNA
A;Residues: 1-607 <BRO>
A;Cross-references: EMBL:X17055; NID:gl386; PIDN:CAA34903.1; PID:gl387
C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
teroid hormones (weak bonds with these hormones promote their transfer across the mem
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; Plasma
F;1-18/Domain: signal sequence #status predicted <SIG>

QY 301 DLPSLAADPVESKDVCKNVAEKDVFLGMFLYIYARRHPDYSVLLRLAKYETTLK 360
 Db 325 DLPLAVDFVEDKEVCKNVEQKDVFLGFLYIYARRHPDYSVLLRLAKYETTLK 384
 QY 361 CAADPHCEYAKVDFEKFPLVEBPQNL 387
 Db 385 CATDDPPACVAFVDFEKFPLVEBPQNL 411

RESULT 4
 ABOS
 serum albumin precursor - horse
 C:Species: Equus caballus (domestic horse)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S34053
 R:Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
 A:Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
 Eur. J. Biochem. 215, 205-212, 1993
 A:Reference number: S34053; MUID:93345495; PMID:8344282
 A:Accession: S34053
 A:Molecule type: mRNA
 A:Residues: 1-607 <HOA>
 A:Cross-references: GB:X74045; NID:G999671; PIDN:CAA52194.1; PID:G999672
 C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membra
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRO>
 F:25-607/Product: serum albumin #status predicted <MAT>
 F:220-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
 F:263/Binding site: bilirubin (lys) #status predicted

Query Match 80.3%; Score 1660.5; DB 1; Length 607;
 Best Local Similarity 77.3%; Pred. No. 4.2e-116;
 Matches 299; Conservative 40; Mismatches 47; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFKDGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
 Db 25 DTHKSEIAHRFNDEGEHFKGLVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES 84
 QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
 Db 85 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 143
 QY 121 DVMTAFHDNEETFLKYLIEIARRHPYFYAPPELLFAKRYKAAFTCCQADKAACLLP 180
 Db 144 DAQCAAFQEDPKFVIGKYLIEIARRHPYFYAPPELLFAKRYKAAFTCCQADKAACLLP 203
 QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
 Db 204 KLDALKEILLSSAKERLKSSQNFGERAVKANSVARLSQKFPKADFAEYVKLVTDLT 263
 QY 241 VITECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 Db 264 VHEKCECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 323
 QY 301 DLPSLAADPVESKDVCKNVAEKDVFLGMFLYIYARRHPDYSVLLRLAKYETTLK 360
 Db 324 DLPALAAADFAEDKEIKYHKKAKADVFLGFLYIYARRHPDYSVLLRLAKYETTLK 383

RESULT 5
 ABOS
 serum albumin precursor - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: S01382; A61006
 R:Weinstock, J.; Baldwin, G.S.
 Nucleic Acids Res. 16, 9045, 1988
 A:Title: Nucleotide sequence of porcine liver albumin.
 A:Reference number: S01382; MUID:89016582; PMID:3174440
 A:Accession: S01382
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-605 <WEI>
 A:Cross-references: EMBL:X12422; NID:G1875; PIDN:CAA30970.1; PID:G833798
 R:Lineback, H.; Sakarya, H.; Chu, W.; Mackinnon, M.
 J. Bone Miner. Res. 4, 235-241, 1989
 A:Title: Serum albumin and its acid hydrolysis peptides dominate preparations of miner
 A:Reference number: A61006; MUID:89268769; PMID:2728927
 A:Accession: A61006
 A:Molecule type: protein
 A:Residues: 23-51,'X',53-54,'XXXGY',146,'E',148,'E',150-151,'XV',155 <LIM>
 A:Experimental source: dental enamel
 A:Note: albumin and other serum proteins are also found in bone
 C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the memba
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; metal binding; plasma
 F:1-16/Domain: signal sequence (fragment) #status predicted <SIG>
 F:17-22/Domain: propeptide #status predicted <PRO>
 F:23-605/Product: serum albumin #status predicted <MAT>
 F:27-199/Domain: serum albumin repeat homology <SA1>
 F:218-391/Domain: serum albumin repeat homology <SA2>
 F:410-589/Domain: serum albumin repeat homology <SA3>
 F:75-84,97-113,112-123,145-190,189-198,221-287,266-274,286-300,299-310,337-382,381-390,
 F:261/Binding site: bilirubin (lys) #status predicted

Query Match 78.5%; Score 1623.5; DB 1; Length 605;
 Best Local Similarity 75.2%; Pred. No. 2.4e-113;
 Matches 291; Conservative 49; Mismatches 46; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFKDGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
 Db 23 DTYKSEIAHRFNDEGEHFKGLVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES 82
 QY 61 NCDKSLHTLFGDKLCTVATLTRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
 Db 83 NCDKSLHTLFGDKLCAIPSLREHYGDLADCKEKEPERNECFLOHKDDNPNLPLVRPEV 141
 QY 121 DVMTAFHDNEETFLKYLIEIARRHPYFYAPPELLFAKRYKAAFTCCQADKAACLLP 180
 Db 142 VALCADFQDEQKFWGKYLIEIARRHPYFYAPPELLYAIYKDFVSECCQADKAACLLP 201
 QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
 Db 202 KIEHLREKVLTSAAKQRLKASIQKFGERAFKANSVARLSQKFPKADFAEYVKLVTDLT 261
 QY 241 VITECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 Db 262 VHEKCECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 321
 QY 301 DLPSLAADPVESKDVCKNVAEKDVFLGMFLYIYARRHPDYSVLLRLAKYETTLK 360
 Db 322 DLNPLEHDFVEDEKVKCNKAEKDVFLGFLYIYARRHPDYSVLLRLAKYETTLK 381
 QY 361 CAADPHCEYAKVDFEKFPLVEBPQNL 387
 Db 382 CAKEDPPACVATVDFKQFLVDEPKNL 408

RESULT 6
 ABOS
 serum albumin precursor [validated] - bovine
 N:Alternate names: 67K protein; prealbumin
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000

F:166-174/Product: kinetensin #status experimental <KIP>
 F:221-394/Domain: serum albumin repeat homology <SA2>
 F:413-592/Domain: serum albumin repeat homology <SA3>
 F:27/Binding site: copper (His) #status predicted
 F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393,4
 F:214/Binding site: pyridoxal phosphate (Lys) (covalent) #status experimental

Query Match 100.0%; Score 2068; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 1.8e-146; Indels 0; Gaps 0;
 Matches 387; Conservative 0; Mismatches 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPEPDHVKLVNEVTEFAKTCVADESAAE 60
 DB 25 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPEPDHVKLVNEVTEFAKTCVADESAAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 144
 QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQQAADKAACLLP 180
 DB 145 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQQAADKAACLLP 204
 QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
 DB 205 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 264
 QY 241 VHTCCCHGDLLECGADDRADLAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 300
 DB 265 VHTCCCHGDLLECGADDRADLAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 324
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
 DB 325 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 384
 QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNL 387
 DB 385 CAADPHCEYAKVDFEFPKPLVEEPQNL 411

RESULT 2
 A47391
 Serum albumin precursor - rhesus macaque
 C:Species: Macaca mulatta (rhesus macaque)
 C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
 C:Accession: A47391
 F:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dwulet, J.; Putnam, F.
 Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
 A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in bilirubin binding
 A:Reference number: A47391; MUID:93211971; PMID:8460152
 A:Contents: B/B homozygote
 A:Accession: A47391
 A:Status: preliminary
 A:Molecule type: mRNA; protein
 A:Residues: 1-600 <WAT>
 A:Cross-references: GB:M90463; NID:g342294; PIDN:AAA36906.1; PID:g342295
 A:Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBI:P128281)
 C:Superfamily: serum albumin; serum albumin repeat homology
 F:21-194/Domain: serum albumin repeat homology <SA1>
 F:213-386/Domain: serum albumin repeat homology <SA2>
 F:405-584/Domain: serum albumin repeat homology <SA3>

Query Match 96.1%; Score 1988; DB 2; Length 600;
 Best Local Similarity 94.6%; Pred. No. 1.6e-140;
 Matches 366; Conservative 13; Mismatches 8; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPEPDHVKLVNEVTEFAKTCVADESAAE 60
 DB 17 DTHKSEVAHRFKDLGEEHFKGLVAVAFSQYLQCCPFBEHVKLVNEVTEFAKTCVADESAAE 76
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 120

Db 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 136
 QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQQAADKAACLLP 180
 Db 137 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQQAADKAACLLP 196
 QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
 Db 197 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 256
 QY 241 VHTCCCHGDLLECGADDRADLAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 300
 Db 257 VHTCCCHGDLLECGADDRADLAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 316
 QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
 Db 317 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 376
 QY 361 CAADPHCEYAKVDFEFPKPLVEEPQNL 387
 Db 377 CAADPHCEYAKVDFEFPKPLVEEPQNL 403

RESULT 3

S57632

serum albumin precursor - cat

C:Species: Felis silvestris catus (domestic cat)

C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999

C:Accession: J04660; S57632

R:Nilgier, C.; Grigioni, F.; Hentges, F.

Gene 169, 235-296, 1996

A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.

A:Reference number: J04660; MUID:96194824; PMID:8647469

A:Accession: J04660

A:Molecule type: mRNA

A:Residues: 1-608 <HI2>

A:Cross-references: EMBL:X84842; NID:9886484; PIDN:CAA59279.1; PID:9886485

A:Experimental source: liver

A:Comment: This protein is the major protein component in plasma. It functions as a mu

ein has 35 conserved cysteine residues

C:Superfamily: serum albumin; serum albumin repeat homology

C:Keywords: liver; plasma

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-24/Domain: propeptide #status predicted <PRP>

F:25-608/Product: serum albumin #status predicted <MAT>

F:29-202/Domain: serum albumin repeat homology <SA1>

F:221-394/Domain: serum albumin repeat homology <SA2>

F:413-592/Domain: serum albumin repeat homology <SA3>

Query Match 83.5%; Score 1727; DB 2; Length 608;

Best Local Similarity 80.6%; Pred. No. 4.6e-121;

Matches 312; Conservative 34; Mismatches 41; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAYLOQCPEPDHVKLVNEVTEFAKTCVADESAAE 60

Db 25 EAHQSEIAHRFNDLGEHFRGLVAVAFSQYLQCCPFBEHVKLVNEVTEFAKTCVADESAAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 120

Db 85 NCEKSLHLLGLDKLCTVASLDRDKYGMADCCCKEKEPERNECFLOHKDDNPGQLVTPEA 144

QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCQQAADKAACLLP 180

Db 145 DAMCTAFHENEQRFGLKLYEIAARRHPYFYAPPELLYAEYKGVFTECEAADKAACLTTP 204

QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240

Db 205 KYDALREKVLASSAKERLKCASLQKFGERAFKAWAVARLSQRPKAEFAEISKLVTDLAK 264

QY 241 VHTCCCHGDLLECGADDRADLAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 300

Db 265 IHKECCHGDLLECGADDRADLAKYICENQDSISKLKECKEPLLEKSHCIAEVENDEMPA 324

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants
R;Menaya, J.; Parrilla, R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A;Reference number: G08292
A;Accession: G01747
A;Status: translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-120, 'G', 122-455 <MEN>
A;Cross-references: EMBL:U22961; NID:9763428; PIDN:AAA64922.1; PID:9763431
R;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
A;Reference number: S55314; MUID:95275251; PMID:7755581
A;Accession: S55314
A;Molecule type: protein
A;Residues: 19-27 <LED>
R;Meloun, B.; Moravek, L.; Kostka, V.
FEBS Lett. 58, 134-137, 1975
A;Title: Complete amino acid sequence of human serum albumin.
A;Reference number: A91420; MUID:76187907; PMID:1225573
A;Accession: A91420
A;Molecule type: protein
A;Residues: 25-117, 'EQ', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-
R;Roehr, U.; Spitteller, G.; Tripiet, D.
Justus Liebig's Ann. Chem. 9, 861-884, 1988
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from ur
A;Reference number: S06422
A;Note: this paper is in German, with an English abstract
A;Accession: S06422
A;Molecule type: protein
R;Finch, J.W.; Crouch, R.K.; Knapp, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A;Title: Mass spectrometric identification of modifications to human serum albumin treat
A;Reference number: S36882; MUID:93384321; PMID:8373198
A;Accession: S36882
A;Molecule type: protein
A;Residues: 45-67, 141-160, 311-337, 469-490, 570-591 <FIN>
R;Kausler, E.; Spitteller, G.
Biol. Chem. Hoppe-Seyler 372, 849-855, 1991
A;Title: Bruchatuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol
A;Reference number: S17599; MUID:92126241; PMID:1772598
A;Accession: S17599
A;Molecule type: protein
A;Residues: 25-54, 354-357, 431-447 <KAU>
A;Note: 49-Leu was also found
R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1694, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: A45800
A;Molecule type: protein
A;Residues: 166-173 <CAR>
R;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
Biochem. Biophys. Res. Commun. 136, 983-988, 1996
A;Title: The amino acid sequence of kinetensin, a novel peptide isolated from pepsin-tre
A;Reference number: A03239; MUID:86242180; PMID:3087352
A;Accession: A03239
A;Molecule type: protein
A;Residues: 166-173, 'L' <MOG>
R;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A;Title: Mutations in genetic variants of human serum albumin found in Italy.
A;Reference number: A38255; MUID:91062352; PMID:2247440
A;Accession: A38255
A;Molecule type: protein
A;Residues: 82-105, 'K', 107-110 <GAL2>
A;Note: this variant is designated albumin Vibo Valentia
A;Accession: A38255
A;Molecule type: protein

A;Residues: 76-83, 'K', 85-106 <GAL3>
A;Note: this variant is designated albumin Torino
R;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin He
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: protein
A;Residues: 255-263, 'B', 265-281 <MIN1>
A;Note: this variant is designated albumin Herborn
R;Minchiotti, L.; Galliano, M.; Stopponi, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta
Biochim. Biophys. Acta 1119, 239-238, 1992
A;Title: Two albumins with identical electrophoretic mobility are produced by diff
A;Reference number: S21078; MUID:92190239; PMID:1347703
A;Accession: S21078
A;Molecule type: protein
A;Residues: 354-356, 'K', 358-378 <MIN2>
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported
R;He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A;Title: Atomic structure and chemistry of human serum albumin.
A;Reference number: A46756; MUID:92334427; PMID:1630489
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
R;Brown, J.R.; Shockley, P.; Behrens, P.Q.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-41
A;Reference number: A94442
A;Contents: annotation; three-dimensional structure and disulfide bonds
R;Saber, M.A.; Stockbauer, P.; Moravek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A;Title: Disulfide bonds in human serum albumin.
A;Reference number: A90930
A;Contents: annotation; disulfide bonds
R;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A;Reference number: A90299; MUID:78186630; PMID:656055
A;Contents: annotation; bilirubin-binding site
R;Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20
A;Title: Serum albumin: conformation and active sites.
A;Reference number: A94408
A;Contents: annotation; binding sites
R;Harper, M.E.; Dugaiczky, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene
A;Reference number: A90028; MUID:83279982; PMID:6192711
A;Contents: annotation; gene position
R;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic ac
A;Reference number: A46755; MUID:76257808; PMID:955075
A;Contents: annotation
A;Note: the nonenzymatic transfer of an acetyl group from aspirin (acetylsalicylic ac
R;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-269, 1992
A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosph
A;Reference number: A56294; MUID:92183981; PMID:1544460
A;Contents: annotation
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in
atase activity
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weat
C;Comment: A large number of variants of human serum albumin have been described.
C;Genetics:
A;Gene: GDB:ALB
A;Cross-references: GDB:118990; OMIM:103600
A;Map position: 4q11-4q13
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyrid
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-609/Product: serum albumin #status experimental <MPT>
F;29-202/Domain: serum albumin repeat homology <SAL>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 47.169 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18_COPY_1_387

Perfect score: 2068

Sequence: 1 DAHKEVAHRFKDLGENFK.....ECYAKVFDEKPLVEEPQNL 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: Piri.*

2: Piri2.*

3: Piri3.*

4: Piri4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2068	100.0	609	1 ABHUS	serum albumin prec
2	1988	96.1	600	2 A47391	serum albumin prec
3	1727	83.5	608	2 S57632	serum albumin prec
4	1660.5	80.3	607	1 ABHOS	serum albumin prec
5	1623.5	78.5	605	1 ABPGS	serum albumin prec
6	1618.5	78.3	607	1 ABPOS	serum albumin prec
7	1613.5	78.0	607	1 ABRS	serum albumin prec
8	1610	77.9	608	1 ABRS	serum albumin prec
9	1585	76.6	609	2 JCS838	albumin - Mongolia
10	1264	61.1	453	2 A05139	serum albumin - mo
11	1038.5	50.2	615	1 ABCHS	serum albumin prec
12	898	43.4	265	2 I46986	albumin - dog (fra
13	766	37.0	607	1 ABXL72	74K albumin precu
14	737.5	35.7	608	1 ABXL68	68K serum albumin
15	724.5	35.0	609	2 JC4358	alpha-fetoprotein
16	721.5	34.9	609	1 PFPU	alpha-fetoprotein
17	720.5	34.8	609	1 PFPU	alpha-fetoprotein
18	691	33.4	599	1 A54906	afamin precursor -
19	656	31.7	608	2 A53195	alpha-fetoprotein
20	616	29.8	605	1 PFMS	alpha-fetoprotein
21	603	29.2	611	1 FPRT	serum albumin prec
22	572.5	27.7	614	2 S59317	serum albumin prec
23	468	22.6	608	1 ABONS1	serum albumin 1 pr
24	465	22.5	608	1 ABONS2	serum albumin 2 pr
25	357	17.3	474	1 VYHD	vitamin D-binding
26	353	17.1	382	2 A37253	serum albumin - bu
27	352	17.0	476	1 VYRD	vitamin D-binding
28	341	16.5	472	1 A35327	vitamin D-binding
29	311	15.0	1423	1 S27941	serum albumin - se

RESULT 1

ABHUS

serum albumin precursor [validated] - human

N:Alternate names: preproalbumin

N:Contains: kinetensin

C:Species: Homo sapiens (man)

C:Date: 29-Jul-1981 #sequence revision 31-Jan-1997 #text change 17-Mar-2000

C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422; S2

R:Law, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houch, C.M.; Najarian, R.C.; Seeb

Nucleic Acids Res. 9, 6103-6114, 1981

A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia col

A:Reference number: A93743; MUID:82081882; PMID:6171778

A:Accession: A93743

A:Molecule type: mRNA

A:Residues: 1-419, 'K', 421-609 <LAW>

A:Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:928591; PIDN:CAA

R:Dugaiczky, A.; Law, S.W.; Demison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A:Reference number: A93936; MUID:82105994; PMID:6275391

A:Accession: A93936

A:Molecule type: mRNA

A:Residues: 1-120, 'G', 122-609 <DUG>

A:Cross-references: EMBL:V00494; NID:928589; PIDN:CAA23753.1; PID:928590

R:Umano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and

A:Reference number: I39427; MUID:86140099; PMID:2419329

A:Accession: I39427

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-26 <URA>

A:Cross-references: GB:M13075; NID:9178330; PIDN:AAAS1688.1; PID:9553173

R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family

A:Reference number: I59286; MUID:94181575; PMID:8134387

A:Accession: I59286

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 282-290, 'KSRPDLQ' <WAT>

A:Cross-references: GB:S69192; NID:9546032; PIDN:AAAB30282.1; PID:9546033

A>Note: This frame-shift variant, designated albumin Roma, leads to analbuminemia

R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl

A:Reference number: I59313; MUID:94294404; PMID:8022807

A:Accession: I59313

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 589-590, 'ALPRVKNLLQVKLP' <WAD>

A:Cross-references: GB:S70799; NID:9547231; PIDN:AAAB31177.1; PID:9547232

cag island protein
cag pathogenicity
mat101-1 protein -
hypothetical prote
serum albumin, mil
major surface glyco
cell surface glyco
phosphoinositide 3
hypothetical prote
glycoprotein anti
embryonic muscle m
hypothetical prote
ENBP1 protein - ba
sperm tail-specifi
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 15

US-10-153-064-125
 ; Sequence 125, Application US/10153064
 ; Patent No. 6663485
 ; GENERAL INFORMATION:
 ; APPLICANT: Bell et al.
 ; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PF556
 ; CURRENT APPLICATION NUMBER: US/10/153,064
 ; CURRENT FILING DATE: 2002-05-24
 ; PRIOR APPLICATION NUMBER: 60/293,212
 ; PRIOR FILING DATE: 2001-05-25
 ; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 125
 ; LENGTH: 677
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-153-064-125

Query Match 100.0%; Score 3103; DB 4; Length 677;
 Best Local Similarity 100.0%; Pred. No. 7.3e-287;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	DAHKSEVAHFRKDLGKGFALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADES	60
Db	93	DAHKSEVAHFRKDLGKGFALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADES	152
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV	120
Db	153	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV	212
Qy	121	DVMCTAHDNEEFLLKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCQADKAACLLP	180
Db	213	DVMCTAHDNEEFLLKYLVEIARRHPYFYAPPELLFPAKYKAAFTCCQADKAACLLP	272
Qy	181	KLDELDRDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEVSGLVTDLT	240
Db	273	KLDELDRDEGKASSAKQRLKASLQKGERAFKAWAVARLSQRPFAEVSGLVTDLT	332
Qy	241	VHTECCHGDLLECADRADLAKYICENQDSISSKLECECKPLLEKSHCIAEVENDMPA	300
Db	333	VHTECCHGDLLECADRADLAKYICENQDSISSKLECECKPLLEKSHCIAEVENDMPA	392
Qy	301	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTTLEK	360
Db	393	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVARRHPDYSVLLRLAKTYETTTLEK	452
Qy	361	CAAADPHECYAKVPDEFKPLVEBPQNLIKQNCLEFQLGKYPQNALVRYTKKVPQVST	420
Db	453	CAAADPHECYAKVPDEFKPLVEBPQNLIKQNCLEFQLGKYPQNALVRYTKKVPQVST	512
Qy	421	PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES	480
Db	513	PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES	572
Qy	481	LVNRRPCFSALEVDETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT	540
Db	573	LVNRRPCFSALEVDETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT	632
Qy	541	KEQLKAVMDDFAAVFVEKCKADDOCKTCFAEEGKKLVAAASQAALGL	585
Db	633	KEQLKAVMDDFAAVFVEKCKADDOCKTCFAEEGKKLVAAASQAALGL	677

Search completed: April 19, 2004, 12:05:17
 Job time : 90.6676 secs

QY 361 CAADPHCEYAKVDFEPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKVQVST 420
DB 432 CAADPHCEYAKVDFEPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKVQVST 491
QY 421 PTLVEVSRNLGKVGSKCKKPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 492 PTLVEVSRNLGKVGSKCKKPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 551
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVVKHKPKAT 540
DB 552 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVVKHKPKAT 611
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
DB 612 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 656

RESULT 13
US-10-153-064-127
; Sequence 127, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 127
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-127

Query Match 100.0%; Score 3103; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 7.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60
DB 92 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 151
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 152 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 211
QY 121 DVMTAFHDNBEETFLKYLIEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 212 DVMTAFHDNBEETFLKYLIEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 271
QY 181 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240
DB 272 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 331
QY 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDMPA 300
DB 332 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDMPA 391
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 360
DB 392 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 451
QY 361 CAADPHCEYAKVDFEPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKVQVST 420
DB 452 CAADPHCEYAKVDFEPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKVQVST 511
QY 421 PTLVEVSRNLGKVGSKCKKPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 512 PTLVEVSRNLGKVGSKCKKPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 571
QY 361 CAADPHCEYAKVDFEPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKVQVST 420
DB 452 CAADPHCEYAKVDFEPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKVQVST 511
QY 421 PTLVEVSRNLGKVGSKCKKPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 512 PTLVEVSRNLGKVGSKCKKPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 571

QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVVKHKPKAT 540
DB 572 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVVKHKPKAT 631
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
DB 632 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 676

RESULT 14
US-10-153-064-129
; Sequence 129, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 129
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-064-129

Query Match 100.0%; Score 3103; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 7.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 60
DB 92 DAHSEVAHRFKDILGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAB 151
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 152 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 211
QY 121 DVMTAFHDNBEETFLKYLIEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 212 DVMTAFHDNBEETFLKYLIEIARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 271
QY 181 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 240
DB 272 KLDELREGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLT 331
QY 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDMPA 300
DB 332 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDMPA 391
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 360
DB 392 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVVLRLAKTYETTTLEKC 451
QY 361 CAADPHCEYAKVDFEPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKVQVST 420
DB 452 CAADPHCEYAKVDFEPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKVQVST 511
QY 421 PTLVEVSRNLGKVGSKCKKPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 512 PTLVEVSRNLGKVGSKCKKPEAKRMPKADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 571
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVVKHKPKAT 540
DB 572 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQATALVELVVKHKPKAT 631
QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
DB 632 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 676

QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 188 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 247
QY 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWARLSORPPKAEFAEVSCLVTDLT 240
DB 248 KLDELDEGKASSAKQRLKASLOKGERAFKAWARLSORPPKAEFAEVSCLVTDLT 307
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
DB 308 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 367
QY 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
DB 368 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 427
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPOVST 420
DB 428 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPOVST 487
QY 421 PTLVEVSRLGKVGSKCKKHPKAEKMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480
DB 488 PTLVEVSRLGKVGSKCKKHPKAEKMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 547
QY 481 LVNRRPCFSALEVEDTVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540
DB 548 LVNRRPCFSALEVEDTVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVHKPKAT 607
QY 541 KEQLKAVMDFFAFAVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 608 KEQLKAVMDFFAFAVEKCKKADDKETCFABEGKKLVAASQAALGL 652

RESULT 11

US-10-153-064-131
; Sequence 131, Application US/10153064
; Patent No. 6663485

; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064

; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 131

; LENGTH: 653

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-153-064-131

Query Match 100.0%; Score 3103; DB 4; Length 653;
Best Local Similarity 100.0%; Pred. No. 6.9e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 69 DAHSEVAHRPKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 128

QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
DB 129 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 188

QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 189 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 248

QY 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWARLSORPPKAEFAEVSCLVTDLT 240
DB 249 KLDELDEGKASSAKQRLKASLOKGERAFKAWARLSORPPKAEFAEVSCLVTDLT 308

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
DB 309 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 368
QY 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
DB 369 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 428
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPOVST 420
DB 429 CAAADPHECYAKVDFEFKPLVEEPQNLIKONCELFQELGEYKFNQALLVRYTKVPOVST 488
QY 421 PTLVEVSRLGKVGSKCKKHPKAEKMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 480
DB 489 PTLVEVSRLGKVGSKCKKHPKAEKMPKCAEDYLSVNLQLCVLHEKTPVSDRVTKCCTES 548
QY 481 LVNRRPCFSALEVEDTVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVHKPKAT 540
DB 549 LVNRRPCFSALEVEDTVVPKEFNAETFTFHADICTLSEKERQIKKQATLVELVHKPKAT 608
QY 541 KEQLKAVMDFFAFAVEKCKKADDKETCFABEGKKLVAASQAALGL 585
DB 609 KEQLKAVMDFFAFAVEKCKKADDKETCFABEGKKLVAASQAALGL 653

RESULT 12

US-10-153-064-130

; Sequence 130, Application US/10153064

; Patent No. 6663485

; GENERAL INFORMATION:

; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064

; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212

; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 130

; LENGTH: 656

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-153-064-130

Query Match 100.0%; Score 3103; DB 4; Length 656;
Best Local Similarity 100.0%; Pred. No. 7e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRPKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 72 DAHSEVAHRPKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 131

QY 61 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
DB 132 NCDSLSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 191

QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 192 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 251

QY 181 KLDELDEGKASSAKQRLKASLOKGERAFKAWARLSORPPKAEFAEVSCLVTDLT 240
DB 252 KLDELDEGKASSAKQRLKASLOKGERAFKAWARLSORPPKAEFAEVSCLVTDLT 311

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
DB 312 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 371

QY 301 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
DB 372 DLPSLAADFVESKDVCKNYAAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 431

QY	1	DAKSEVAHRFKDLG	GENFKALVLI	FAQY	LOCCP	PEDHVKLV	NEVTEFAKTCV	ADESAE	60
DB	67	DAKSEVAHRFKDLG	GENFKALVLI	FAQY	LOCCP	PEDHVKLV	NEVTEFAKTCV	ADESAE	126
QY	61	NCDSLHTLFGDKLCTV	ATRETYGEMADCCAKQ	PERNECFLOHKD	NDNPNL	PRLV	PREV	120	
DB	127	NCDSLHTLFGDKLCTV	ATRETYGEMADCCAKQ	PERNECFLOHKD	NDNPNL	PRLV	PREV	186	
QY	121	DVMTAFHNEETFLK	LYIETARRHPYI	YAPPELLFFAKRY	KAAATTECCQA	ADKAA	CLLP	180	
DB	187	DVMTAFHNEETFLK	LYIETARRHPYI	YAPPELLFFAKRY	KAAATTECCQA	ADKAA	CLLP	246	
QY	181	KLDELDEGKASSAK	ORLKASLOKFG	ERAFKAWAVARLS	QRPFPKAE	FAEVSKLV	TDLT	240	
DB	247	KLDELDEGKASSAK	ORLKASLOKFG	ERAFKAWAVARLS	QRPFPKAE	FAEVSKLV	TDLT	306	
QY	241	VHTECHGDLLECAD	DRADLAKYICENQ	DISISKLEKCEK	PILLESK	HCIAE	VNDENMPA	300	
DB	307	VHTECHGDLLECAD	DRADLAKYICENQ	DISISKLEKCEK	PILLESK	HCIAE	VNDENMPA	366	
QY	301	DLPSLAADFVESK	VDKCNVAEAKD	VLGMFLVEYARR	HPDYSV	LVLLRL	LAKEYITTELC	360	
DB	367	DLPSLAADFVESK	VDKCNVAEAKD	VLGMFLVEYARR	HPDYSV	LVLLRL	LAKEYITTELC	426	
QY	361	CAAADPHECVAK	YDFEFKPLVEBP	QNLIKONCEL	PEQLGEYFQ	NALLVRYTK	VPQVST	420	
DB	427	CAAADPHECVAK	YDFEFKPLVEBP	QNLIKONCEL	PEQLGEYFQ	NALLVRYTK	VPQVST	486	
QY	421	PTLVEVSRNLGK	VGSKCKKHP	EAKRMPCAEDY	LSVVLNQL	CVLHEKTP	PSDRVTKCCTES	480	
DB	487	PTLVEVSRNLGK	VGSKCKKHP	EAKRMPCAEDY	LSVVLNQL	CVLHEKTP	PSDRVTKCCTES	546	
QY	481	LVNRRPFCFSALE	VDYETVVPKEF	NAETFTTHADICTL	SEKERQIKK	QTALVELV	VGHKPKAT	540	
DB	547	LVNRRPFCFSALE	VDYETVVPKEF	NAETFTTHADICTL	SEKERQIKK	QTALVELV	VGHKPKAT	606	
QY	541	KEQLKAVMDDFA	AFVVEKCKC	ADDKETCFABEG	KKLVAASQA	ALGL	585		
DB	607	KEQLKAVMDDFA	AFVVEKCKC	ADDKETCFABEG	KKLVAASQA	ALGL	651		
<p>RESULT 10</p> <p>US-10-153-064-132</p> <p>; Sequence 132, Application US/10153064</p> <p>; Patent No. 6663485</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Bell et al.</p> <p>; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins</p> <p>; FILE REFERENCE: PF556</p> <p>; CURRENT APPLICATION NUMBER: US/10/153,064</p> <p>; PRIOR FILING DATE: 2002-05-24</p> <p>; PRIOR FILING DATE: 2001-05-25</p> <p>; NUMBER OF SEQ ID NOS: 137</p> <p>; SOFTWARE: Patent in version 3.1</p> <p>; SEQ ID NO 132</p> <p>; LENGTH: 652</p> <p>; TYPE: PRT</p> <p>; ORGANISM: Homo sapiens</p> <p>US-10-153-064-132</p> <p>Query Match 100.0%; Score 3103; DB 4; Length 652;</p> <p>Best Local Similarity 100.0%; Pred. NO. 6.9e-287;</p> <p>Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;</p>									
QY	1	DAKSEVAHRFKDLG	GENFKALVLI <td>FAQY</td> <td>LOCCP</td> <td>PEDHVKLV</td> <td>NEVTEFAKTCV</td> <td>ADESAE</td> <td>60</td>	FAQY	LOCCP	PEDHVKLV	NEVTEFAKTCV	ADESAE	60
DB	68	DAKSEVAHRFKDLG	GENFKALVLI <td>FAQY</td> <td>LOCCP</td> <td>PEDHVKLV</td> <td>NEVTEFAKTCV</td> <td>ADESAE</td> <td>127</td>	FAQY	LOCCP	PEDHVKLV	NEVTEFAKTCV	ADESAE	127
QY	61	NCDSLHTLFGDKLCTV	ATRETYGEMADCCAKQ	PERNECFLOHKD	NDNPNL	PRLV	PREV	120	
DB	128	NCDSLHTLFGDKLCTV	ATRETYGEMADCCAKQ	PERNECFLOHKD	NDNPNL	PRLV	PREV	187	

QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
DB 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585

RESULT 5

US-10-153-064-7
; Sequence 7, Application US/10153064
; Patent No. 6663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match 100.0%; Score 3103; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.2e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRPKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60
DB 25 DAHSEVAHRPKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELREGKASSAKQRLKASLOKFGERAPKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 205 KLDELREGKASSAKQRLKASLOKFGERAPKAWAVARLSORFPKAEFAEVS KLVTDLTK 264
QY 241 VHTCCCHGDLLECADDDRADLAKYICENODSIS SKLKECCCKP LLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADDDRADLAKYICENODSIS SKLKECCCKP LLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVL LRLAKTYETTLEKC 360
DB 324 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVL LRLAKTYETTLEKC 384
QY 361 CAADDPHECYAKVDFEKP LVEEPQNLIKQNCSELFQGLGEYKFNALLVRYTKKVPQVST 420
DB 385 CAADDPHECYAKVDFEKP LVEEPQNLIKQNCSELFQGLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKPEAKRMPCAEDYLSVNLQNL CVLHEKTPVSDRVT KCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKPEAKRMPCAEDYLSVNLQNL CVLHEKTPVSDRVT KCCTES 504
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
DB 565 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 609

RESULT 7

US-08-797-689-2
; Sequence 2, Application US/08797689

RESULT 6

US-09-976-594-977
; Sequence 777, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furtress, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 977
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 088957CD1
US-09-976-594-977

Query Match 100.0%; Score 3103; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.2e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRPKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESA 60
DB 25 DAHSEVAHRPKDLGEENFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 145 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204
QY 181 KLDELREGKASSAKQRLKASLOKFGERAPKAWAVARLSORFPKAEFAEVS KLVTDLTK 240
DB 205 KLDELREGKASSAKQRLKASLOKFGERAPKAWAVARLSORFPKAEFAEVS KLVTDLTK 264
QY 241 VHTCCCHGDLLECADDDRADLAKYICENODSIS SKLKECCCKP LLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECADDDRADLAKYICENODSIS SKLKECCCKP LLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVL LRLAKTYETTLEKC 360
DB 325 DLPSLAADFVBSKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVVL LRLAKTYETTLEKC 384
QY 361 CAADDPHECYAKVDFEKP LVEEPQNLIKQNCSELFQGLGEYKFNALLVRYTKKVPQVST 420
DB 385 CAADDPHECYAKVDFEKP LVEEPQNLIKQNCSELFQGLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKKPEAKRMPCAEDYLSVNLQNL CVLHEKTPVSDRVT KCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKKPEAKRMPCAEDYLSVNLQNL CVLHEKTPVSDRVT KCCTES 504
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
DB 505 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 564
QY 541 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
DB 565 KEQLKAVMDPFAAFVEKCKKADDKETCFABEGKLVAAASQAALGL 609

QY 541 KEQLKAVMDFAAFVEKCKCKKADKCTCFABEGKLVAAASQAALGL 585
DB 541 KEQLKAVMDFAAFVEKCKCKKADKCTCFABEGKLVAAASQAALGL 585

RESULT 3

US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonmenschin, Carlos
; APPLICANT: Soco, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match 100.0%; Score 3103; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.9e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGKLVIAFAQYLOCCPPEDHVKLVNEVTEFAKTCVADESAB 60
DB 1 DAHSEVAHRFKDLGKLVIAFAQYLOCCPPEDHVKLVNEVTEFAKTCVADESAB 60
QY 61 NCDSLSHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
DB 61 NCDSLSHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
QY 121 DVNCTAFHNDNEETPLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVNCTAFHNDNEETPLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASQKGERAFKAWAVARLSORPPKAEFAVSKLVTDLT 240
DB 181 KLDELDEGKASSAKQRLKASQKGERAFKAWAVARLSORPPKAEFAVSKLVTDLT 240
QY 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLKECCCKPLEKSHGICAEVNDMPA 300
DB 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLKECCCKPLEKSHGICAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLRKAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLRKAKTYETTLK 360

QY 361 CAADPHCEYAKVDFDFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADPHCEYAKVDFDFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKYGSKCKKHPKAPKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
DB 421 PTLVEVSRNLGKYGSKCKKHPKAPKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKQTALVELVKHKPKAT 540
QY 541 KEQLKAVMDFAAFVEKCKCKKADKCTCFABEGKLVAAASQAALGL 585
DB 541 KEQLKAVMDFAAFVEKCKCKKADKCTCFABEGKLVAAASQAALGL 585

RESULT 4

US-10-153-064-5
; Sequence 5, Application US/10153064
; Patent No. 8663485
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 3103; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.9e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGKLVIAFAQYLOCCPPEDHVKLVNEVTEFAKTCVADESAB 60
DB 1 DAHSEVAHRFKDLGKLVIAFAQYLOCCPPEDHVKLVNEVTEFAKTCVADESAB 60
QY 61 NCDSLSHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
DB 61 NCDSLSHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKKDDNPNLRLVRPEV 120
QY 121 DVNCTAFHNDNEETPLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVNCTAFHNDNEETPLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASQKGERAFKAWAVARLSORPPKAEFAVSKLVTDLT 240
DB 181 KLDELDEGKASSAKQRLKASQKGERAFKAWAVARLSORPPKAEFAVSKLVTDLT 240
QY 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLKECCCKPLEKSHGICAEVNDMPA 300
DB 241 VHTCCGHDLLCADDRAADLAKYICENQDSISSKLKECCCKPLEKSHGICAEVNDMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLRKAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLLRKAKTYETTLK 360
QY 361 CAADPHCEYAKVDFDFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADPHCEYAKVDFDFKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKYGSKCKKHPKAPKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
DB 421 PTLVEVSRNLGKYGSKCKKHPKAPKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480

MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: Region
 LOCATION: 369..419
 OTHER INFORMATION: /note= "Alternative C-termini of
 OTHER INFORMATION: HSA(1-n)"
 FEATURE:
 NAME/KEY: Region
 LOCATION: 1..585
 OTHER INFORMATION: /note= "Amino acid sequence of
 OTHER INFORMATION: natural HSA"
 US-08-153-799-14

Query Match 100.0%; Score 3103; DB 1; Length 585;
 Best Local Similarity 100.0%; Pred. No. 5.9e-287;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFLOHKDDNPNLRLVPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFLOHKDDNPNLRLVPEV 120

QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTETTTLEKC 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTETTTLEKC 360

QY 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420

QY 421 PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
 DB 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540

RESULT 2
 US-08-702-572-2
 ; Sequence 2, Application US/08702572
 ; Patent No. 5965386
 ; GENERAL INFORMATION:
 ; APPLICANT: Kerry-Williams, Sean M
 ; APPLICANT: Gilbert, Sarah C
 ; TITLE OF INVENTION: Yeast Strains and Modified Albumins
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Centeon L.L.C.
 ; STREET: 1020 First Avenue

CITY: King of Prussia
 STATE: Pennsylvania
 COUNTRY: USA
 ZIP: 19406-1310
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: Microsoft Word 6.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/702,572
 FILING DATE: 11-NOV-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 95/23857
 FILING DATE: 1-MAR-1995
 APPLICATION NUMBER: GB 9404270.2
 FILING DATE: 5-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Naomi Biswas
 REGISTRATION NUMBER: 38,384
 REFERENCE/DOCKET NUMBER: CE0114 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610/878/4294
 TELEFAX: 610/878/4221
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 585 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-702-572-2

Query Match 100.0%; Score 3103; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 5.9e-287;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFLOHKDDNPNLRLVPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEBNECFLOHKDDNPNLRLVPEV 120

QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELDRDEGKASSAKQRLKASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTETTTLEKC 360
 DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLRLAKTETTTLEKC 360

QY 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADAPHECYAKVDFEFPKPLVEEPQNLIKONCELFEQLGEYKFNALLVRYTKKVPQVST 420

QY 421 PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
 DB 481 LVNRRPCFSALEVDVETVYVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:40:29 ; Search time 89.6676 Seconds

(without alignments)
336.813 Million cell updates/sec

Title: US-09-832-929-18

Perfect score: 3103

Sequence: 1 DAKSEVAHFQDLGLENPK.....TCFAEKGKLVAAQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/protdata/2/iaa/5A COMB.pep.*
- 2: /cgn2_6/protdata/2/iaa/5B COMB.pep.*
- 3: /cgn2_6/protdata/2/iaa/6A COMB.pep.*
- 4: /cgn2_6/protdata/2/iaa/6B COMB.pep.*
- 5: /cgn2_6/protdata/2/iaa/PCUS COMB.pep.*
- 6: /cgn2_6/protdata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	585	1	US-08-153-799-14
2	3103	100.0	585	2	US-08-702-572-2
3	3103	100.0	585	3	US-08-769-746-2
4	3103	100.0	585	4	US-10-153-064-5
5	3103	100.0	609	4	US-10-153-064-7
6	3103	100.0	609	4	US-09-576-594-977
7	3103	100.0	610	2	US-08-797-689-2
8	3103	100.0	610	4	US-09-984-186-2
9	3103	100.0	651	4	US-10-153-064-133
10	3103	100.0	652	4	US-10-153-064-132
11	3103	100.0	653	4	US-10-153-064-131
12	3103	100.0	656	4	US-10-153-064-130
13	3103	100.0	676	4	US-10-153-064-127
14	3103	100.0	676	4	US-10-153-064-129
15	3103	100.0	677	4	US-10-153-064-125
16	3103	100.0	680	4	US-10-153-064-123
17	3103	100.0	783	1	US-08-256-938-2
18	3103	100.0	787	1	US-08-256-938-4
19	3103	100.0	787	2	US-08-797-689-16
20	3103	100.0	787	4	US-09-984-186-16
21	3099	99.9	609	1	US-08-222-619-3
22	3099	99.9	609	1	US-08-433-037-4
23	3099	99.9	609	4	US-08-897-956A-2
24	3099	99.9	609	5	PCT-US95-04075-3
25	3095	99.7	978	4	US-08-897-956A-3
26	3093	99.7	585	1	US-08-448-196A-3
27	3093	99.7	585	2	US-08-984-176-1

28	3092.5	99.7	652	4	US-10-153-064-96
29	3092.5	99.7	652	4	US-10-153-064-99
30	3092.5	99.7	652	4	US-10-153-064-105
31	3092.5	99.7	660	4	US-10-153-064-90
32	3092.5	99.7	660	4	US-10-153-064-93
33	3092.5	99.7	676	4	US-10-153-064-95
34	3092.5	99.7	676	4	US-10-153-064-98
35	3092.5	99.7	676	4	US-10-153-064-104
36	3092.5	99.7	684	4	US-10-153-064-92
37	3092.5	99.7	1184	4	US-10-153-064-89
38	3092.5	99.5	668	4	US-10-153-064-102
39	3086.5	99.5	692	4	US-10-153-064-101
40	2458.5	79.2	583	1	US-08-448-196A-5
41	2450.5	79.0	583	1	US-08-448-196A-4
42	2432.5	78.4	583	1	US-08-448-196A-6
43	2426	78.2	584	1	US-08-448-196A-7
44	2389	77.0	582	1	US-08-134-638-1
45	1249.5	40.3	609	1	US-08-222-619-4

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murkay Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/153.799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

Sequence 96, Appl
Sequence 99, Appl
Sequence 105, Appl
Sequence 90, Appl
Sequence 93, Appl
Sequence 95, Appl
Sequence 98, Appl
Sequence 104, Appl
Sequence 92, Appl
Sequence 89, Appl
Sequence 102, Appl
Sequence 101, Appl
Sequence 5, Appl
Sequence 4, Appl
Sequence 6, Appl
Sequence 7, Appl
Sequence 1, Appl

; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match 100.0%; Score 38; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 15

US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match 100.0%; Score 38; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

Search completed: April 19, 2004, 12:54:58
Job time : 4.36842 secs

```
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18

Query Match      100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 11
US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Beltzer, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLyS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-932-322-445

Query Match      100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 12
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 13
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 38; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 14
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
```

/ / TELECOMMUNICATION INFORMATION:
/ / TELEPHONE: (415) 705-8410
/ / TELEFAX: (415) 397-8338
/ / INFORMATION FOR SEQ ID NO: 2:
/ / SEQUENCE CHARACTERISTICS:
/ / LENGTH: 585 amino acids
/ / TYPE: amino acid
/ / TOPOLOGY: linear
/ / MOLECULE TYPE: protein
/ / SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match 100.0%; Score 38; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 7
US-09-932-613-445
Sequence 445, Application US/09932613
Publication No. US20030091565A1
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
APPLICANT: Beltzer, James P.
APPLICANT: Potter, M. Daniel
APPLICANT: Fleming, Tony J.
TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US
CURRENT APPLICATION NUMBER: US/09/932,613
CURRENT FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 458
SOFTWARE: Patent in version 3.1
SEQ ID NO 445
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-932-613-445

Query Match 100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 8
US-09-984-010-26
Sequence 26, Application US/09984010
Publication No. US20030104578A1
GENERAL INFORMATION:
APPLICANT: Ballance, David James
TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
AND SERUM ALBUMIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSER: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
STREET: 1300 I Street, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3315
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)

/ / CURRENT APPLICATION DATA:
/ / APPLICATION NUMBER: US/09/984,010
/ / FILING DATE: 21-May-2002
/ / PRIOR APPLICATION DATA:
/ / APPLICATION NUMBER: US 09/091,873
/ / FILING DATE: 25-JUN-1998
/ / APPLICATION NUMBER: PCT/CB96/03164
/ / FILING DATE: 19-DEC-1996
/ / INFORMATION FOR SEQ ID NO: 26:
/ / SEQUENCE CHARACTERISTICS:
/ / LENGTH: 585 amino acids
/ / TYPE: amino acid
/ / STRANDEDNESS: <Unknown>
/ / TOPOLOGY: linear
/ / MOLECULE TYPE: protein
/ / HYPOCHETICAL: NO
/ / ANTI-SENSE: NO
/ / SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 9
US-09-833-041-18
Sequence 18, Application US/09833041
Publication No. US20030125247A1
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A.
APPLICANT: Haseltine, William A.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF545
CURRENT APPLICATION NUMBER: US/09/833,041
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 79
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 18
LENGTH: 585
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-833-041-18

Query Match 100.0%; Score 38; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 54 VADESAEN 61

RESULT 10
US-09-833-117-18
Sequence 18, Application US/09833117
Publication No. US20030171267A1
GENERAL INFORMATION:
APPLICANT: Rosen, Craig A.
APPLICANT: Sadeghi, Homayoon
APPLICANT: Prior, Christopher P.
APPLICANT: Turner, Andrew J.
TITLE OF INVENTION: Albumin Fusion Proteins

```

; TITLE OF INVENTION: CHIMERIC POLYPEPTIDES OF SERUM ALBUMIN AND USES RELATED THERETO
; FILE REFERENCE: GPCI-P01-109
; CURRENT APPLICATION NUMBER: US/09/768,183
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 09/764918
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 09/619285
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/144534
; PRIOR FILING DATE: 1999-07-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-768-183-6

Query Match      100.0%; Score 38; DB 12; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      1 VADESAEN 8
Db      4 VADESAEN 11

RESULT 3
US-10-074-956-24
; Sequence 24, Application US/10074956
; Publication No. US2002019332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-074-956-24

Query Match      100.0%; Score 38; DB 13; Length 195;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      1 VADESAEN 8
Db      54 VADESAEN 61

RESULT 4
US-10-074-956-27
; Sequence 27, Application US/10074956
; Publication No. US2002019332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-074-956-27

Query Match      100.0%; Score 38; DB 13; Length 241;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      1 VADESAEN 8
Db      78 VADESAEN 85

RESULT 5
US-10-074-956-28
; Sequence 28, Application US/10074956
; Publication No. US2002019332A1
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: METHODS OF TREATING BLADDER DISORDERS
; FILE REFERENCE: 08191-022001
; CURRENT APPLICATION NUMBER: US/10/074,956
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/268,175
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-074-956-28

Query Match      100.0%; Score 38; DB 13; Length 268;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      1 VADESAEN 8
Db      78 VADESAEN 85

RESULT 6
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnenschein, Carlos
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 3.36842 Seconds
(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_54_61
Perfect score: 38
Sequence: 1 VADESAEN 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCR_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/FCRUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	100.0	14	12	US-09-764-918-6
2	38	100.0	14	12	US-09-768-183-6
3	38	100.0	195	13	US-10-074-956-24
4	38	100.0	241	13	US-10-074-956-27
5	38	100.0	268	13	US-10-074-956-28
6	38	100.0	585	9	US-09-929-552-2
7	38	100.0	585	10	US-09-932-613-445
8	38	100.0	585	10	US-09-984-010-26
9	38	100.0	585	10	US-09-833-041-18
10	38	100.0	585	10	US-09-833-117-18
11	38	100.0	585	10	US-09-932-322-445
12	38	100.0	585	10	US-09-832-501-18
13	38	100.0	585	11	US-09-833-118-18
14	38	100.0	585	11	US-09-833-245-18
15	38	100.0	585	12	US-10-424-999-11

16	38	100.0	585	12	US-10-425-000-31	Sequence 31, Appl
17	38	100.0	585	12	US-10-433-108-34	Sequence 34, Appl
18	38	100.0	585	13	US-10-153-064-5	Sequence 5, Appl
19	38	100.0	585	14	US-10-153-604A-5	Sequence 5, Appl
20	38	100.0	585	14	US-10-319-263-1	Sequence 1, Appl
21	38	100.0	585	14	US-10-319-263-2	Sequence 2, Appl
22	38	100.0	585	14	US-10-414-489-1	Sequence 1, Appl
23	38	100.0	585	14	US-10-414-469-2	Sequence 2, Appl
24	38	100.0	585	14	US-10-413-831-1	Sequence 1, Appl
25	38	100.0	585	14	US-10-413-831-2	Sequence 2, Appl
26	38	100.0	585	15	US-10-413-832-1	Sequence 1, Appl
27	38	100.0	585	15	US-10-413-832-2	Sequence 2, Appl
28	38	100.0	585	15	US-10-414-386-1	Sequence 1, Appl
29	38	100.0	585	15	US-10-414-386-2	Sequence 2, Appl
30	38	100.0	585	15	US-10-233-675A-11	Sequence 11, Appl
31	38	100.0	585	15	US-10-462-262-26	Sequence 26, Appl
32	38	100.0	604	10	US-09-984-010-7	Sequence 7, Appl
33	38	100.0	609	10	US-09-919-039-370	Sequence 370, App
34	38	100.0	609	12	US-10-609-346-12	Sequence 12, Appl
35	38	100.0	609	13	US-10-153-064-7	Sequence 7, Appl
36	38	100.0	609	14	US-10-153-604A-7	Sequence 7, Appl
37	38	100.0	609	14	US-10-365-623-23	Sequence 23, Appl
38	38	100.0	610	9	US-09-984-186-2	Sequence 2, Appl
39	38	100.0	610	14	US-10-237-667-2	Sequence 2, Appl
40	38	100.0	610	14	US-10-237-708-2	Sequence 2, Appl
41	38	100.0	610	14	US-10-237-866-2	Sequence 2, Appl
42	38	100.0	610	14	US-10-237-871-2	Sequence 2, Appl
43	38	100.0	610	14	US-10-237-871-2	Sequence 2, Appl
44	38	100.0	616	12	US-10-433-108-13	Sequence 13, Appl
45	38	100.0	624	12	US-10-433-108-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-764-918-6
; Sequence 6, Application US/09764918
; Publication No. US20010056075A1
; GENERAL INFORMATION:
; APPLICANT: Gyuris, J.
; TITLE OF INVENTION: CHIMERIC POLYPEPTIDES OF SERUM ALBUMIN AND USES RELATED THERETO
; FILE REFERENCE: GPC1-P02-109
; CURRENT APPLICATION NUMBER: US/09/764,918
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 09/619285
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-764-918-6

Query Match 100.0%; Score 38; DB 12; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VADESAEN 8
Db 4 VADESAEN 11

RESULT 2
US-09-768-183-6
; Sequence 6, Application US/09768183
; Publication No. US20020048571A1
; GENERAL INFORMATION:
; APPLICANT: Gyuris, J.
; APPLICANT: Lamphere, L.
; APPLICANT: Morris, A.

```

PR 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX
XX PA (ZYCO-) ZYCO INC.
XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX WPI; 2002-195801/25.
DR
XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
XX Example 2; Page 48; 89pp; English.
PS
XX The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a peptide described in the exemplification of the
CC invention
XX
XX Sequence 236 AA;
SQ
Query Match 100.0%; Score 38; DB 5; Length 236;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
Db 78 VADESAEN 85
|||||
|||||

Search completed: April 19, 2004, 11:51:16
Job time : 5.70545 secs

```

XX Yeast codon-biased recombinant HSA protein fragment HSA-I.
 DE Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
 XX overlapping oligonucleotide; expression vector.
 KW Homo sapiens.
 OS Synthetic.
 OS CN1239103-A.
 PN 22-DEC-1999.
 XX 17-JUN-1998; 98CN-00102506.
 XX 17-JUN-1998; 98CN-00102506.
 XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.
 PA Li S, Lu D;
 XX WPI; 2000-351198/31.
 DR N-PSDB; AAA10092.
 XX Process for preparing recombinant human serum albumin comprising yeast
 PT biased sex codons - uses a recombinant DNA technique.
 XX Example 1; Fig 3; 44pp; Chinese.
 XX The method relates to a method of recombinantly producing human serum
 CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
 CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as
 CC three synthetic fragments (AAA10092-AAA10094) joined by recombinant DNA
 CC technology. Each HSA fragment was synthesised from overlapping
 CC oligonucleotide fragments that were extended. This sequence represents
 CC the sequence of the HSA fragment HSA-I encoded by the human gene with a
 CC yeast codon bias. The invention also covers a recombinant expression
 CC vector, yeast host cells carrying the recombinant expression vector and
 CC the process for producing human serum albumin in the yeast host cell,
 CC especially in secretory mode
 XX Sequence 204 AA;
 SQ

Query Match 100.0%; Score 38; DB 3; Length 204;
 Best Local Similarity 100.0%; Pred. No. 6.2; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;
 QY 1 VADESAEN 8
 DB 61 VADESAEN 68

RESULT 14
 AAU29874
 ID AAU29874 standard; protein; 214 AA.
 XX AAU29874;
 AC AAU29874;
 XX 18-DEC-2001 (first entry)
 DT Novel human secreted protein #365.
 DE Human; vaccination; gene therapy; nutritional supplement;
 XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX Homo sapiens.
 OS WO200179449-A2.
 PN 25-OCT-2001.
 XX 16-APR-2001; 2001WO-US008656.

XX 18-APR-2000; 2000US-00552929.
 PR 26-JAN-2001; 2001US-00770160.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-611725/70.
 DR Nucleic acids encoding a range of human polypeptides, useful in genetic
 XX vaccination, testing and therapy.
 PT Claim 20; Page 205; 765pp; English.
 XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX Sequence 214 AA;
 SQ

Query Match 100.0%; Score 38; DB 4; Length 214;
 Best Local Similarity 100.0%; Pred. No. 6.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VADESAEN 8
 DB 90 VADESAEN 97

RESULT 15
 AAU17051
 ID AAU17051 standard; protein; 236 AA.
 XX AAU17051;
 AC AAU17051;
 XX 29-MAY-2002 (first entry)
 DT Human albumin-thrombin-alpha-MSH SEQ ID NO: 82.
 DE Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
 XX alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
 KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
 KW immunosuppressive; anti-inflammatory; antirheumatic; antiarthritic;
 KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
 KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
 KW diabetes; uveitis; coeliac disease.
 XX Homo sapiens.
 OS WO200206316-A2.
 PN 24-JAN-2002.
 XX 16-JUL-2001; 2001WO-US022263.
 XX 14-JUL-2000; 2000US-0218381P.
 PR 18-AUG-2000; 2000US-0226382P.
 XX 06-OCT-2000; 2000US-0238380P.

RESULT 11
 AAO17048
 ID AAO17048 standard; protein; 195 AA.
 XX AC AAO17048;
 XX DT 29-MAY-2002 (first entry)
 XX DE Human serum albumin (1-195) SEQ ID NO: 56.
 XX KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
 XX KW Alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
 XX KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
 XX KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
 XX KW antiasnatic; antibacterial; dermatological; antipsoriatic;
 XX KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
 XX KW diabetes; uveitis; coeliac disease.
 XX OS Homo sapiens.
 XX PN W0200206316-A2.
 XX PD 24-JAN-2002.
 XX PF 16-JUL-2001; 2001WO-US022263.
 XX PR 14-JUL-2000; 2000US-0218381P.
 XX PR 18-AUG-2000; 2000US-0226382P.
 XX PR 06-OCT-2000; 2000US-0328380P.
 XX PR 29-DEC-2000; 2000US-0238764P.
 XX PR 14-JUN-2001; 2001US-0298317P.
 XX PA (ZYCO-) ZYCOS INC.
 XX PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
 XX WPI; 2002-195801/25.
 XX PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
 XX PT stimulating hormone concatamer or its analog, for treating inflammatory
 XX PT or autoimmune disorders.
 XX PS Example 2; Page 46; 89pp; English.
 XX CC The present invention relates to a nucleic acid comprising a sequence
 XX CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
 XX CC hormone (MSH) concatamer. The sequences are useful for treating an
 XX CC individual suffering from, or at risk of, a disorder of the immune system
 XX CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
 XX CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
 XX CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
 XX CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
 XX CC present sequence is a protein described in the exemplification of the
 XX CC invention
 XX SQ Sequence 195 AA;
 Query Match 100.0%; Score 38; DB 5; Length 195;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VADESAEN 8
 |||||
 DB 54 VADESAEN 61
 RESULT 12
 ABU10022
 ID ABU10022 standard; protein; 195 AA.
 XX AC ABU10022;
 XX DT 28-JUL-2000 (first entry)

DT 31-JUL-2003 (first entry)
 XX DE Human serum albumin residues 1-195.
 XX KW Bladder disorder; cytostatic; antiinflammatory; immune response;
 KW un-methylated CpG sequence; alpha-MSH; melanocortin receptor;
 KW bladder cancer; tumour; interstitial cystitis; inflammation;
 KW alpha-MSH concatamer; melanocyte stimulating hormone; human;
 KW serum albumin.
 XX OS Homo sapiens.
 XX PN US2002193332-A1.
 XX PD 19-DEC-2002.
 XX PF 12-FEB-2002; 2002US-00074956.
 XX PR 12-FEB-2001; 2001US-0268175P.
 XX PA (HEDL/) HEDLEY M L.
 XX PI Hedley ML;
 XX WPI; 2003-447327/42.
 XX PT Modulating immune responses in a mammal with a bladder disorder e.g.
 PT bladder cancer, by administering nucleic acids comprising un-methylated
 PT CpG sequences, nucleic acids encoding alpha-MSH, or alpha-MSH peptides to
 PT the mammal.
 XX PS Example 2; Page 9; 17pp; English.
 XX CC The invention describes a method of modulating an immune response in a
 CC mammal, comprising identifying a mammal that has or is at risk for having
 CC a bladder disorder, and administering: (a) an isolated nucleic acid (N1)
 CC comprising an un-methylated CpG sequence to the mammal; (b) an isolated
 CC nucleic acid (N2) comprising sequence encoding alpha-MSH to the mammal;
 CC or (c) a peptide that binds to a melanocortin receptor to the mammal. The
 CC method is useful for modulating immune responses in a mammal having a
 CC bladder disorder, where administration of (N1) results in an amelioration
 CC of one or more symptoms of the disorder. Preferably, the method is useful
 CC for modulating immune response in a mammal having bladder cancer (where
 CC administration of (N1) results in a decrease in tumour size or activity),
 CC or for modulating immune response in a mammal having interstitial
 CC cystitis (where administration of (N1) results in a modulation of the
 CC immune response from Th2 response to a Th1 response). The method is also
 CC useful for modulating immune response in a mammal having bladder disorder
 CC that is characterised by inflammation which is associated with symptoms
 CC of interstitial cystitis or associated with a disruption of the integrity
 CC of the bladder lining. This is the amino acid sequence of human serum
 CC albumin residues 1-195 that can be used in the creation of melanocyte
 CC stimulating hormone (alpha-MSH) concatamers resulting in secretion of the
 CC fusion protein when expressed in mammalian cells
 XX SQ Sequence 195 AA;
 Query Match 100.0%; Score 38; DB 7; Length 195;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VADESAEN 8
 |||||
 DB 54 VADESAEN 61
 RESULT 13
 AAY83947
 ID AAY83947 standard; protein; 204 AA.
 XX AC AAY83947;
 XX DT 28-JUL-2000 (first entry)

CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 137 AA;

Query Match 100.0%; Score 38; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
| | | | |
Db 83 VADESAEN 90

RESULT 9
AAU33073
ID AAU33073 standard; protein; 156 AA.
XX
AC AAU33073;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3564.
XX
Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
XX
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 704; 765pp; English.
XX

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
XX
SQ Sequence 156 AA;

QY 1 VADESAEN 8
| | | | |
Db 83 VADESAEN 90

Query Match 100.0%; Score 38; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
| | | | |
Db 9 VADESAEN 16

RESULT 10
AAU33271
ID AAU33271 standard; protein; 156 AA.
XX
AC AAU33271;
XX

DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3762.
XX
Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PF 16-APR-2001; 2001WO-US008656.
XX
PR 18-APR-2000; 2000US-00552929.
XX
PR 26-JAN-2001; 2001US-00770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
PS Claim 20; Page 751; 765pp; English.
XX

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration; immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU29510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention
XX
XX
SQ Sequence 156 AA;

Query Match 100.0%; Score 38; DB 4; Length 156;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VADESAEN 8
| | | | |
Db 9 VADESAEN 16

```

PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
PI
DR WPI; 2001-514838/56.
DR N-PSDB; AA182560.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX
XX Claim 20; SEQ ID NO 16521; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 127 AA;
SQ
Query Match 100.0%; Score 38; DB 4; Length 127;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VADESAEN 8
DB 38 VADESAEN 45
|||||
RESULT 7
AAO09589
ID AAO09589 standard; protein; 132 AA.
AC AAO09589;
XX
XX 06-NOV-2001 (first entry)
DT
DE Human polypeptide SEQ ID NO 23481.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX 07-SEP-2001.
PD
XX 26-FEB-2001; 2001WO-US004927.
PF
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
DR
DR N-PSDB; AA189520.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX
XX Claim 20; SEQ ID NO 16521; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 132 AA;
SQ
Query Match 100.0%; Score 38; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VADESAEN 8
DB 82 VADESAEN 89
|||||
RESULT 8
AAO12088
ID AAO12088 standard; protein; 137 AA.
XX
XX AAO12088;
AC
XX 06-NOV-2001 (first entry)
DT
XX Human polypeptide SEQ ID NO 25980.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX 07-SEP-2001.
PD
XX 26-FEB-2001; 2001WO-US004927.
PF
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
DR
DR N-PSDB; AA192019.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX
XX Claim 20; SEQ ID NO 25980; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity, tissue growth factor activity, haematopoiesis regulating
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 132 AA;
SQ
Query Match 100.0%; Score 38; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VADESAEN 8
DB 82 VADESAEN 89
|||||

```

```

DT 06-NOV-2001 (first entry)
DE Human polypeptide SEQ ID NO 18339.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX N-PSDB; AAI84378.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 18339; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and/or
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 120 AA;
XX
XX Query Match 100.0%; Score 38; DB 4; Length 120;
XX Best Local Similarity 100.0%; Pred. No. 3.4;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 VADESAEN 8
XX |||||
XX 18 VADESAEN 25
XX
XX Db
XX
XX RESULT 5
XX AAO04415
XX ID AAO04415 standard; protein; 124 AA.
XX
XX AC AAO04415;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 18307.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX
OS

```

```

XX
PN WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX N-PSDB; AAI84346.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 18307; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and/or
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 124 AA;
XX
XX Query Match 100.0%; Score 38; DB 4; Length 124;
XX Best Local Similarity 100.0%; Pred. No. 3.6;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 VADESAEN 8
XX |||||
XX 76 VADESAEN 83
XX
XX Db
XX
XX RESULT 6
XX AAO02629
XX ID AAO02629 standard; protein; 127 AA.
XX
XX AC AAO02629;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 16521.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX

```


CC three synthetic fragments (AA10092-AL0094) joined by recombinant DNA
CC technology. Each HSA fragment was synthesised from overlapping
CC oligonucleotide fragments that were extended. This sequence represents
CC the complete sequence of the HSA encoded by the human gene with a yeast
CC codon bias. The invention also covers a recombinant expression vector,
CC yeast host cells carrying the recombinant expression vector and the
CC process for producing human serum albumin in the yeast host cell,
CC especially in secretory mode
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAFHFKDGLGSENFKALVLIAPAFYQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAFHFKDGLGSENFKALVLIAPAFYQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVPRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVPRPEV 120
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCSLQKFGGERAFKAWAVARLSORFFKAEFAEYKLVTDLT 240
DB 181 KLDELDEGKASSAKORLKCSLQKFGGERAFKAWAVARLSORFFKAEFAEYKLVTDLT 240
QY 241 VHTECHGDLLECCADDRADLAKYICENQDSISSKLECEKPLEKSHCIAEVENDEMPA 300
DB 241 VHTECHGDLLECCADDRADLAKYICENQDSISSKLECEKPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEAARRHPDYSVLLLRKAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEAARRHPDYSVLLLRKAKTYETTTLEKC 360
QY 361 CAADPHECYAKVDFDEKPLVEEPQNL 387
DB 361 CAADPHECYAKVDFDEKPLVEEPQNL 387

RESULT 15

AAMS2567
ID AAMS2567 standard; protein; 585 AA.

XX AC AAMS2567;

XX DT 05-FEB-2002 (first entry)

XX DE Mature human serum albumin.

XX KW Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiant;
KW neotropic; neuroprotective; gene therapy; immune disorder; wound healing;
KW hyperproliferative disorder; renal disorder; cardiovascular disorder;
KW respiratory disorder; neurological disease; endocrine disorder;
KW reproductive system disorder; infectious disease;
KW gastrointestinal disorder.

XX OS Homo sapiens.

XX PN WO200179444-A2.

XX PD 25-OCT-2001.

XX PF 12-APR-2001; 2001WO-US012013.

XX PR 12-APR-2000; 2000US-0229359P.

XX PR 25-APR-2000; 2000US-0199384P.

XX PR 21-DEC-2000; 2000US-0256931P.

PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Haseltine WA;

XX WI; 2001-616755/71.

DR N-PSDB; ABA03057.

XX PT Albumin fusion proteins comprising a therapeutic protein and albumin,

PT useful in the treating immune system disorders (e.g. transplant

PT rejection), blood related disorders (e.g. myocardial infarction) and

XX PS hyperproliferative disorders.

XX Claim 1; Fig 15; 606pp; English.

CC The present invention relates to albumin fusion proteins, which comprise
CC a therapeutic protein and albumin. The present sequence is the protein
CC sequence for mature human serum albumin (HA), which was used to generate
CC the fusion proteins of the present invention. The albumin fusion proteins
CC are useful in the treatment, prevention, diagnosis, and/or detection of
CC diseases/disorders such as immune system disorders (e.g. transplant
CC rejection), blood related disorders (e.g. myocardial infarction),
CC hyperproliferative disorders (e.g. childhood acute myeloid leukemia),
CC renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
CC arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),
CC neurological diseases (e.g. Alzheimer's disease), endocrine disorders
CC (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),
CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
CC irritable bowel syndrome) and wound healing

XX SQ Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAFHFKDGLGSENFKALVLIAPAFYQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAFHFKDGLGSENFKALVLIAPAFYQYLOCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVPRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVPRPEV 120
QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCSLQKFGGERAFKAWAVARLSORFFKAEFAEYKLVTDLT 240
DB 181 KLDELDEGKASSAKORLKCSLQKFGGERAFKAWAVARLSORFFKAEFAEYKLVTDLT 240
QY 241 VHTECHGDLLECCADDRADLAKYICENQDSISSKLECEKPLEKSHCIAEVENDEMPA 300
DB 241 VHTECHGDLLECCADDRADLAKYICENQDSISSKLECEKPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEAARRHPDYSVLLLRKAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEAARRHPDYSVLLLRKAKTYETTTLEKC 360
QY 361 CAADPHECYAKVDFDEKPLVEEPQNL 387
DB 361 CAADPHECYAKVDFDEKPLVEEPQNL 387

Search completed: April 19, 2004, 11:51:15
Job time : 228.626 secs

QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECEKPELLEKSHCHIAEVENDEMPA 300
DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECEKPELLEKSHCHIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVYARRHPDYSVLLRLAKTYETTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVYARRHPDYSVLLRLAKTYETTLEKC 360
QY 361 CAADPHECYAKVDFEKPVLVEPQNL 387
DB 361 CAADPHECYAKVDFEKPVLVEPQNL 387

RESULT 13
AAY84873
ID AAY84873 standard; protein; 585 AA.
XX AAY84873;
XX
DT 08-AUG-2000 (first entry)
DE Amino acid sequence of a human albumin protein.
XX Human; albumin; ischemic state; serum protein; metal ion salt;
KW perioperative ischemia; ischemia; myocardial infarction;
KW progressive coronary artery disease.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Modified-site 1 /note= "optionally acetylated, and claimed under claim
FT 56"
FT
XX
XX WO200020840-A1.
XX
XX 13-APR-2000.
XX
XX 01-OCT-1999; 99WO-US022905.
XX
XX 02-OCT-1998; 98US-00165581.
XX 02-OCT-1998; 98US-00165926.
XX 02-OCT-1998; 98US-0102738P.
XX 11-JAN-1999; 99US-0115392P.
XX
XX (ISCH-) ISCHEMIA TECHNOLOGIES INC.
XX
XX Bar-Or D, Lau E, Winkler JV;
XX
XX WPI; 2000-303843/26.
XX
XX New method for the continuous detection of ischemic states comprises
FT detecting and quantifying the existence of an alteration of the serum
FT protein albumin.
XX
XX Disclosure; Page 97-100; 105pp; English.

XX The present sequence represents human albumin protein. The specification
CC describes a method for the continuous detection of ischemic states. The
CC method comprises detecting and quantifying the existence of an alteration
CC of the serum protein albumin. The method comprises contacting a
CC biological sample containing albumin from the patient with an excess
CC quantity of a metal ion salt, where the metal ion binds to the N-terminus
CC of naturally occurring human albumin, to form a mixture containing bound
CC metal ions and unbound metal ions, and then determining the amount of
CC metal ions bound to the albumin N-terminus. The amount of bound metal
CC ions is correlated to a known value to determine the occurrence or non-
CC occurrence of an ischemic event. The methods are useful for detection of
CC ischemic states. The methods are also useful for distinguishing
CC perioperative ischemia from ischemia caused by, amongst other things,
CC myocardial infarctions and progressive coronary artery disease
XX
XX Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191; Indels 0; Gaps 0;
Matches 387; Conservative 0; Mismatches 0;
QY 1 DAHSEVAHRFKDLGBENFKALVLIATAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAHRFKDLGBENFKALVLIATAQYLOQCPEFDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQBERNECTLOHKDDNPNIPLRVPRPV 120
DB 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKQBERNECTLOHKDDNPNIPLRVPRPV 120
QY 121 DVMTAFHDNEETFLKKLYEYARRHPYFVAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEYARRHPYFVAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFVSKLVTDLTJK 240
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEFVSKLVTDLTJK 240
QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECEKPELLEKSHCHIAEVENDEMPA 300
DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLECEKPELLEKSHCHIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVYARRHPDYSVLLRLAKTYETTLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEVYARRHPDYSVLLRLAKTYETTLEKC 360
QY 361 CAADPHECYAKVDFEKPVLVEPQNL 387
DB 361 CAADPHECYAKVDFEKPVLVEPQNL 387

RESULT 14
AAY83946
ID AAY83946 standard; protein; 585 AA.
XX AAY83946;
XX
DT 28-JUL-2000 (first entry)
DT
XX Yeast codon-biased recombinant human serum albumin protein.
DE Yeast codon-biased recombinant human serum albumin protein.
XX Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
KW overlapping oligonucleotide; expression vector.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX CN1239103-A.
XX
XX 22-DEC-1999.
XX
XX 17-JUN-1998; 98CN-00102506.
XX
XX 17-JUN-1998; 98CN-00102506.
XX
XX (HAIJ-) HAIJI BIOENGINEERING CO LTD.
XX
XX Li S, Lu D;
XX WPI; 2000-351198/31.
XX N-PSDB; AAA10091.
XX
XX Process for preparing recombinant human serum albumin comprising yeast
FT biased sex codons - uses a recombinant DNA technique.
XX
XX Disclosure; Fig 1; 44pp; Chinese.
XX
XX The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
CC a yeast codon bias. The complete HSA gene (AAA10091) was generated as

KW Serum albumin-growth hormone fusion protein; growth hormone;
KW Down's syndrome.
XX
XX Unidentified.
XX KR99076789-A.
XX
XX 15-OCT-1999.
XX
XX 25-JUN-1998; 98KR-00704914.
XX
XX 30-DEC-1995; 95GB-00026733.
XX 19-DEC-1996; 96WO-GB003164.
XX
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
XX Ballance DJ;
XX
XX WPI; 1997-363680/33.
XX N-PSDB; AAK99568.
XX
XX Serum albumin-growth hormone fusion protein - useful to treat growth
XX hormone related diseases, e.g. Down's syndrome.
XX
XX Disclosure; Fig 6; 21pp; Korean.
XX
XX The invention relates to a serum albumin-growth hormone fusion protein -
XX useful to treat growth hormone related diseases such as Down's syndrome.
XX This sequence represents a HSA protein related to the serum albumin-
XX growth hormone protein of the invention
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 2068; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAFQYLOQCPEFEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHSEVAHRFKDLGSENFKALVLIAPAFQYLOQCPEFEDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVWCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
DB 121 DVWCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGRAFAMAVARLSQRPFAKAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKCSLQKFGRAFAMAVARLSQRPFAKAEVSKLVTDLTK 240
QY 241 VITECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDEMPA 300
DB 241 VITECHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDEMPA 300
QY 301 DLPSLAADPVEKQVCKVAEAKDVLGFLNFIYEYARRHPDYSVLLLLRLAKTYETTLK 360
DB 301 DLPSLAADPVEKQVCKVAEAKDVLGFLNFIYEYARRHPDYSVLLLLRLAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
DB 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
RESULT 12
ID AAW59841 standard; protein; 585 AA.
XX
XX AAW59841;
XX AC
XX
XX 20-NOV-1998 (first entry)
DT

XX Mature protein of human serum albumin (HSA).
DE
XX
XX Protein expression; monocytledon plant cell;
KW glycosylated alpha 1-antitrypsin; AAT; glycosylated antithrombin III;
KW AAT; human serum albumin; HSA; subtilisin BPN'; treatment; emphysema;
KW antithrombotic; blood replacement.
XX
XX Homo sapiens.
XX
XX WO9836085-A1.
XX
XX 20-AUG-1998.
XX
XX 13-FEB-1998; 98WO-US003068.
XX
XX 13-FEB-1997; 97US-0037991P.
XX 13-FEB-1997; 97US-0038168P.
XX 13-FEB-1997; 97US-0038169P.
XX 13-FEB-1997; 97US-0038170P.
XX
XX (PHYT-) APPLIED PHYTOLOGICS INC.
XX
XX Sutliff TD, Rodriguez RL;
XX
XX WPI; 1998-467179/40.
XX N-PSDB; AAV41728.
XX
XX Expressing mature, glycosylated proteins in monocytledonous plant cells
XX - from chimeric gene including signal peptide sequence, specifically
XX therapeutic agents and industrial enzymes.
XX
XX Disclosure; Page 31-32; 53pp; English.
XX
XX The present sequence represents the mature protein of human serum albumin
XX (HSA). The protein is used to exemplify the invention. The specification
XX describes a method for producing mature heterologous protein in
XX monocytledonous plant cells. The method comprises transforming the cells
XX with a chimeric gene comprising a monocytledon transcription regulator,
XX inducible either during seed maturation or by adding/removing a small
XX molecule, DNA encoding the heterologous protein, and DNA encoding a
XX signal peptide, with the signal peptide causing secretion of the protein
XX from the cell. Proteins expressed in this manner include mature
XX glycosylated alpha 1-antitrypsin (AAT) with a glycosylation pattern that
XX significantly increases its serum half-life, mature glycosylated
XX antithrombin III (ATIII), mature human serum albumin (HSA) having the
XX native folding pattern as shown by bilirubin-binding characteristics, or
XX mature active subtilisin BPN'. These proteins are useful therapeutically
XX (e.g. AAT for treating emphysema, ATIII as antithrombotic and HSA as
XX blood replacement) or as industrial enzymes (BPN' is used in detergents)
XX
XX Sequence 585 AA;
SQ
Query Match 100.0%; Score 2068; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGSENFKALVLIAPAFQYLOQCPEFEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHSEVAHRFKDLGSENFKALVLIAPAFQYLOQCPEFEDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVWCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
DB 121 DVWCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGRAFAMAVARLSQRPFAKAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKCSLQKFGRAFAMAVARLSQRPFAKAEVSKLVTDLTK 240

PN JP04211375-A.
 XX 03-AUG-1992.
 XX 05-FEB-1991; 91JP-00014600.
 XX 05-FEB-1990; 90JP-00025682.
 XX (AJIN) AJINOMOTO KK.
 XX WPI; 1992-304940/37.
 XX N-PSDB; AAQ27813.
 XX Synthetic gene for prepn. of human serum albumin - comprises synthetic
 PT DNA contg. gene coding the albumin using coding in Escherichia coli.
 XX Fig 1; Page 13; 37pp; Japanese.
 XX The protein sequence was deduced from the synthetic DNA sequence encoding
 CC human serum albumin which was prepd. by ligating eight synthetic HSA gene
 CC fragments
 XX Sequence 585 AA;
 Query Match 100.0%; Score 2068; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1.1e-191;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 Db 1 DAHSEVAHRFKDGLGENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEERNECFLOHKDDNPNLRLVRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
 Db 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
 QY 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 Db 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTTLEKC 360
 Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTTLEKC 360
 QY 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387
 Db 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387
 RESULT 10
 AAR80301
 ID AAR80301 standard; protein; 585 AA.
 XX AAR80301;
 XX 25-MAR-2003 (revised)
 DT 17-JAN-1996 (first entry)
 XX Human serum albumin.
 XX Serum albumin; HSA; aspartyl protease-3; Yap3p; Saccharomyces cerevisiae.
 XX Homo sapiens.
 OS Homo sapiens.
 XX W09523857-A1.
 FN

XX 08-SEP-1995.
 XX 01-MAR-1995; 95WC-GB0000434.
 XX 05-MAR-1994; 94GB-00004270.
 XX (DELZ) DELTA BIOTECHNOLOGY LTD.
 XX Kerrywilliams SM, Gilbert SC;
 XX WPI; 1995-320572/41.
 XX N-PSDB; AAQ98695.
 XX Yeast with reduced levels of aspartyl protease 3 proteolytic activity -
 PT used to secrete human albumin without prodn. of the 45 kD fragment.
 XX Example 1; Page 26-28; 50pp; English.
 XX The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected
 CC to site-directed mutagenesis to investigate the role of endoproteases in
 CC the generation of a 45 kDa albumin fragment obtd. when the cDNA is
 CC expressed in S. cerevisiae. Mutations were: R410A; L407A; L408V; V409A;
 CC and R410A, K413Q, K414Q. The latter set of mutations, especially,
 CC improved stability of HSA to yeast Yap3p proteolytic cleavage, allowing
 CC increased prodn. of recombinant HSA. (Updated on 25-MAR-2003 to correct
 CC PI field.)
 XX Sequence 585 AA;
 Query Match 100.0%; Score 2068; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1.1e-191;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHRFKDGLGENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 Db 1 DAHSEVAHRFKDGLGENFKALVLIAPAYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEERNECFLOHKDDNPNLRLVRPEV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPEERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
 Db 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
 QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
 QY 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 Db 241 VHTCCCHGDLLECADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTTLEKC 360
 Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTTLEKC 360
 QY 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387
 Db 361 CAADPHECYAKVDFEFPKPLVEEPQNL 387

RESULT 11
 AAO20111
 ID AAO20111 standard; protein; 585 AA.
 XX AAO20111;
 XX 06-AUG-2002 (first entry)
 DT HSA protein sequence related to the growth hormone protein.
 XX

```

DR N-PSDB; AAQ04719.
XX
XX Human serum albumin prepn. by yeast host - by culturing transformed
PT plasmid yeast to produce serum, and removing it.
XX
XX Disclosure; Page ?; -pp; Japanese.
PS
XX Mature HSA-A may be produced using the sequence incorporated into a
CC plasmid vector with suitable controllers, and transferred to a yeast
CC expression system
XX
XX Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYIQQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYIQQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLTG 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLTG 240
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 8
AAR08457
ID AAR08457 standard; protein; 585 AA.
XX
XX AAR08457;
AC
XX
XX 25-MAR-2003 (revised)
DT 16-APR-1991 (first entry)
XX
XX Human serum albumin.
DE
XX HSA; folding; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Region 1..303
FT /label= B
FT Region 123..585
FT /label= C
FT Region 123..303
FT /label= A
XX
XX JF02227079-A.
XX
XX 10-SEP-1990.
XX

```

```

PF 25-AUG-1989; 89JP-00217540.
XX
XX 06-OCT-1988; 88JP-00250926.
XX
XX (TOFU) TONEN CORP.
PA
XX WPI; 1990-317325/42.
DR N-PSDB; AAQ06099.
XX
XX New human serum albumin fragments - used to bond to medicines and for
PT stable folding of protein(s).
XX
XX Claim 1; Fig 8; 24pp; Japanese.
PS
XX Fragments A-C of HSA are expressed as fusion proteins with the signal
CC peptide of E. coli alkaline phosphatase. The fragments are selected for
CC their specific properties. The C-terminal truncated fragment, B, does not
CC bind long-chain fatty acids but does bind to various medicines at the
CC central region. The N-terminal truncated fragment, C, has good stability
CC in protein folding. The central segment, A, has characteristics of both B
CC and C. See also AAQ06096-Q06098. (Updated on 25-MAR-2003 to correct PD
CC field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-
XX 2003 to correct PR field.)
SQ Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYIQQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYIQQCPFDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLTG 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLTG 240
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVLLLRLLAKTYETTLK 360
Qy 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 9
AAR26362
ID AAR26362 standard; protein; 585 AA.
XX
XX AAR26362;
AC
XX 11-FEB-1993 (first entry)
DT
XX Synthetic HSA protein.
DE
XX Human serum albumin; transformants; recombinant.
XX
XX Synthetic.
XX

```

CC which the codons which are most frequently used by the chosen non-human
CC host were selected. In this case, it is yeast cells (L120; Leu2-3; 112,
CC His 3-11, 15). The synthetic HSA gene was assembled from 24
CC oligonucleotide blocks. HSA is used in therapy for the treatment of
CC hypovolaemia, shock and hypalbuminaemia. It is also used as an additive
CC in perfusion liq. for extracorporeal circulation and as an experimental
CC antigen. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-
XX -2003 to correct PI field.)
SQ Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKD:GEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKD:GEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Qy 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLECCCKPILKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLECCCKPILKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 360
Qy 361 CAADPHCEYAKVDFDEPKPLVEEPQNL 387
Db 361 CAADPHCEYAKVDFDEPKPLVEEPQNL 387

RESULT 6
AAP90388
ID AAP90388 standard; protein; 585 AA.
XX AAP90388;
AC
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX Mature human serum albumin polypeptide.
DE
XX Human serum albumin; mature protein; new polypeptides; plasma expanders.
KW
XX Homo sapiens: (Human).
OS
XX EF322094-A.
PN
XX 28-JUN-1989.
PD
XX 25-OCT-1988; 88EP-00310000.
PF
XX 30-OCT-1987; 87GB-00025529.
PR
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
PA
XX Ballance DJ, Hinchliffe E, Gelsow MJ, Senior PJ;
PI WPI; 1989-186464/26.
XX
DR

DR N-PSDB; AAN90128.
XX New N-terminal fragments of human serum albumin - esp. useful as blood
PT plasma expanders.
XX
XX Disclosure; Fig 2; 20pp; English.
XX Mature protein of human serum albumin (see corresp. AAN90128). Used to
CC make new N-terminal fragments which are used as plasma expanders, or as
CC substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX
SQ Sequence 585 AA;

Query Match 100.0%; Score 2068; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKD:GEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKD:GEENFKALVLIAPAOYLQOCPEEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Qy 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLECCCKPILKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLECCCKPILKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYFYARRHPDYSVLLLRLLAKTYETTTLEKC 360
Qy 361 CAADPHCEYAKVDFDEPKPLVEEPQNL 387
Db 361 CAADPHCEYAKVDFDEPKPLVEEPQNL 387

RESULT 7
AAR05318
ID AAR05318 standard; protein; 585 AA.
XX AAR05318;
AC
DT 08-OCT-1990 (first entry)
XX Human serum albumin gene product.
DE
XX Human serum albumin; HSA-A; yeast; ds.
KW
XX Homo sapiens.
OS
XX JP02117384-A.
PN
XX 01-MAY-1990.
PD
XX 26-OCT-1988; 88JP-00268302.
PF
XX 26-OCT-1988; 88JP-00268302.
PR
XX (TOFU) TOA NENRYO KOGYO KK.
PA
XX WPI; 1990-176228/23.
XX
DR

	Matches	387;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	DAHSEVAHRFDLGEENFKALVLIAPAQYLOQCPEEDHVKLVNEVTEFAKTCVADSEAE	60							
Db	1	DAHSEVAHRFDLGEENFKALVLIAPAQYLOQCPEEDHVKLVNEVTEFAKTCVADSEAE	60							
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVRPEV	120							
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVRPEV	120							
Qy	121	DVMCTAFHNDNEETFLKKLYIETARRHPYFYAPELLFPAKRYKAAPTECCQADKAACLLP	180							
Db	121	DVMCTAFHNDNEETFLKKLYIETARRHPYFYAPELLFPAKRYKAAPTECCQADKAACLLP	180							
Qy	181	KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEVSCLVTDLT	240							
Db	181	KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEVSCLVTDLT	240							
Qy	241	VHTECHGDLLECCADRADLAKYICENQDSISSKLECEKPLEKSHCIAEVENDEMPA	300							
Db	241	VHTECHGDLLECCADRADLAKYICENQDSISSKLECEKPLEKSHCIAEVENDEMPA	300							
Qy	301	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLK	360							
Db	301	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLK	360							
Qy	361	CAAADPHECYAKVDFEFKPLVEEPQNL	387							
Db	361	CAAADPHECYAKVDFEFKPLVEEPQNL	387							

RESULT 4
AAP90392
ID AAP90392 standard; protein; 407 AA.
XX
AC AAP90392;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
XX N-terminal human serum albumin.
DE
XX Human serum albumin; mature protein; new polypeptides; plasma expanders.
XX Homo sapiens; (Human).
XX
XX EP322094-A.
XX
XX 28-JUN-1989.
XX
XX 25-OCT-1988; 88EP-00310000.
XX
XX 30-OCT-1987; 87GB-00025529.
XX
XX (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
XX Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX WPI; 1989-186464/26.
XX
XX New N-terminal fragments of human serum albumin - esp. useful as blood
XX plasma expanders.
XX
XX Claim 2; Page 9; 20pp; English.
XX
XX N-terminal portion of human serum albumin. Used to make plasma expanders,
XX or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25
XX -MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
XX field)
XX
XX Sequence 407 AA;

	Query Match	100.0%;	Score	2068;	DB	1;	Length	407;
	Best Local Similarity	100.0%;	Pred. No.	6.9e-192;				
	Matches	387;	Conservative	0;	Mismatches	0;	Indels	0;
Qy	1	DAHSEVAHRFDLGEENFKALVLIAPAQYLOQCPEEDHVKLVNEVTEFAKTCVADSEAE	60					
Db	1	DAHSEVAHRFDLGEENFKALVLIAPAQYLOQCPEEDHVKLVNEVTEFAKTCVADSEAE	60					
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVRPEV	120					
Db	61	NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPRLVRPEV	120					
Qy	121	DVMCTAFHNDNEETFLKKLYIETARRHPYFYAPELLFPAKRYKAAPTECCQADKAACLLP	180					
Db	121	DVMCTAFHNDNEETFLKKLYIETARRHPYFYAPELLFPAKRYKAAPTECCQADKAACLLP	180					
Qy	181	KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEVSCLVTDLT	240					
Db	181	KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFAEVSCLVTDLT	240					
Qy	241	VHTECHGDLLECCADRADLAKYICENQDSISSKLECEKPLEKSHCIAEVENDEMPA	300					
Db	241	VHTECHGDLLECCADRADLAKYICENQDSISSKLECEKPLEKSHCIAEVENDEMPA	300					
Qy	301	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLK	360					
Db	301	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVVLLLRLLAKTYETTLK	360					
Qy	361	CAAADPHECYAKVDFEFKPLVEEPQNL	387					
Db	361	CAAADPHECYAKVDFEFKPLVEEPQNL	387					

RESULT 5
AAP93344
ID AAP93344 standard; protein; 585 AA.
XX
AC AAP93344;
XX
DT 25-MAR-2003 (revised)
DT 23-JUN-1990 (first entry)
XX
XX Sequence of mature human serum albumin (HSA) as encoded by artificial
XX Gene.
XX Mature human serum albumin; artificial gene; oligonucleotide block;
XX hypobolaemia; shock; hypoalbuminaemia.
XX Homo sapiens.
XX
XX EP308381-A.
XX
XX 22-MAR-1989.
XX
XX 13-SEP-1988; 88EP-00850299.
XX
XX 14-SEP-1987; 87SE-00003539.
XX
XX (SKAN-) SKANDIGEN AB.
XX (MAGY) MTA SZEGEDI BIOLÓG KOZPONTI.
XX (VEPE-) VEPEX CONTRACTOR LTD.
XX
XX Aberg B, Simoncsits A, Kalan M, Csperpan I, Bajszar G;
XX WPI; 1989-087749/12.
XX N-PSDB; AAN90997.
XX
XX Artificial gene coding for authentic human serum albumin - constructed on
XX the basis of codons most frequently used by chosen non-human host.
XX
XX Disclosure; pp. 11-16; 121pp; English.
XX
XX The synthetic gene was constructed by designing a nucleotide sequence in

Db 1 DAHKSEVAHRFKDLGEBNFKALVLIATAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAB 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120
QY 121 DVMCTAFHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
Db 121 DVMCTAFHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
QY 241 VITECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VITECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFVEYARRHPDYVWLLLLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFVEYARRHPDYVWLLLLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 2
AAP90390
ID AAP90390 standard; protein; 389 AA.
XX AC AAP90390;
XX DT 24-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 01-NOV-1989 (first entry)
XX DE N-terminal human serum albumin.
XX KW N-terminal portion of human serum albumin; plasma expanders.
XX OS Homo sapiens; (Human).
XX PN EP322094-A.
XX PD 28-JUN-1989.
XX PF 25-OCT-1988; 88EP-00310000.
XX PR 30-OCT-1987; 87GB-00025529.
XX PA (DELZ) DELTA BIOTECHNOLOGY LTD.
XX PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX DR WPI; 1989-186464/26.
XX PT New N-terminal fragments of human serum albumin - esp. useful as blood
XX PS plasma expanders.
XX PS Claim 2; Page 9; 20pp; English.
XX CC N-terminal portion of human serum albumin. Used to make plasma expanders,
XX CC or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25
XX CC -MAR-2003 to correct FA field.) (Updated on 24-OCT-2003 to standardise OS
XX CC field)
XX SQ Sequence 389 AA;
Query Match 100.0%; Score 2068; DB 1; Length 389;
Best Local Similarity 100.0%; Pred. No. 6.4e-192;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEBNFKALVLIATAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAB 60
Db 1 DAHKSEVAHRFKDLGEBNFKALVLIATAQYLOQCFFEDHVKLVNEVTEFAKTCVADESAB 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVRPEV 120
QY 121 DVMCTAFHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
Db 121 DVMCTAFHNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
QY 241 VITECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 241 VITECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFVEYARRHPDYVWLLLLAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFVEYARRHPDYVWLLLLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387
Db 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 3
AAP90391
ID AAP90391 standard; protein; 390 AA.
XX AC AAP90391;
XX DT 24-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 01-NOV-1989 (first entry)
XX DE N-terminal human serum albumin.
XX KW N-terminal portion of human serum albumin; plasma expanders.
XX OS Homo sapiens; (Human).
XX PN EP322094-A.
XX PD 28-JUN-1989.
XX PF 25-OCT-1988; 88EP-00310000.
XX PR 30-OCT-1987; 87GB-00025529.
XX PA (DELZ) DELTA BIOTECHNOLOGY LTD.
XX PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX DR WPI; 1989-186464/26.
XX PT New N-terminal fragments of human serum albumin - esp. useful as blood
XX PS plasma expanders.
XX PS Claim 2; Page 9; 20pp; English.
XX CC N-terminal portion of human serum albumin. Used to make new N-terminal
XX CC fragments which are used as plasma expanders, or as substitutes for HSA
XX CC or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct FA
XX CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX SQ Sequence 390 AA;
Query Match 100.0%; Score 2068; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 6.5e-192;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 227.626 Seconds
(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18_COPY_1_387

Perfect score: 2068

Sequence: 1 DAKHSEVAHRFKDLEENFK.....ECYAKVFDFKPLVEEPQNL 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:**

1: Geneseq1980s:**

2: Geneseq1990s:**

3: Geneseq2000s:**

4: Geneseq2001s:**

5: Geneseq2002s:**

6: Geneseq2003as:**

7: Geneseq2003bs:**

8: Geneseq2004s:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2068	100.0	388	1 AAP90389	Aap90389 N-termina
2	2068	100.0	389	1 AAP90390	Aap90390 N-termina
3	2068	100.0	390	1 AAP90391	Aap90391 N-termina
4	2068	100.0	407	1 AAP90392	Aap90392 N-termina
5	2068	100.0	585	1 AAP93344	Aap93344 Sequence
6	2068	100.0	585	1 AAP90388	Aap90388 Mature hu
7	2068	100.0	585	2 AAR05318	Aar05318 Human ser
8	2068	100.0	585	2 AAR08457	Aar08457 Human ser
9	2068	100.0	585	2 AAR26362	Aar26362 Synthetic
10	2068	100.0	585	2 AAR80301	Aar80301 Human ser
11	2068	100.0	585	2 AAO20111	Aao20111 HSA prote
12	2068	100.0	585	2 AAW59841	Aaw59841 Mature pr
13	2068	100.0	585	3 AAY84873	Aay84873 Amino aci
14	2068	100.0	585	3 AAY83946	Aay83946 Yeast cod
15	2068	100.0	585	4 AAM52567	Aam52567 Mature hu
16	2068	100.0	585	4 AAE12403	Aae12403 Human alb
17	2068	100.0	585	4 AAE13129	Aae13129 Human alb
18	2068	100.0	585	4 AAE13399	Aae13399 Human alb
19	2068	100.0	585	4 ABB79006	Abb79006 Human mat
20	2068	100.0	585	4 AAE08578	Aae08578 Human ser
21	2068	100.0	585	5 AAU75220	Aau75220 Mature fo
22	2068	100.0	585	5 ABJ00986	Abj00986 B lymphoc
23	2068	100.0	585	5 ABG63321	Abg63321 Human ser
24	2068	100.0	585	5 ABG33847	Abg33847 Human B L
25	2068	100.0	585	5 ABG71291	Abg71291 Glycosyla

26	2068	100.0	585	6 ABR55695	AbR55695 Human alb
27	2068	100.0	585	6 AAE30936	Aae30936 Human alb
28	2068	100.0	585	7 ABR42606	AbR42606 Human ser
29	2068	100.0	585	7 ADC16767	AdC16767 Human ser
30	2068	100.0	585	7 ADC16794	AdC16794 Mutant hu
31	2068	100.0	585	7 ADC16795	AdC16795 Mutant hu
32	2068	100.0	585	7 ADD06469	Add06469 Human ser
33	2068	100.0	585	7 ADD32014	Add32014 Heterolog
34	2068	100.0	585	7 ADD68016	Add68016 Mature fo
35	2068	100.0	608	2 AAR05784	Aar05784 Fragment
36	2068	100.0	608	2 AAR96229	Aar96229 Human ser
37	2068	100.0	609	1 AAP30089	Aap30089 Sequence
38	2068	100.0	609	2 AAR94572	Aar94572 Cancer me
39	2068	100.0	609	2 AAR88913	Aar88913 Human ser
40	2068	100.0	609	2 AAR96232	Aar96232 Human ser
41	2068	100.0	609	2 AAW48095	Aaw48095 Human ser
42	2068	100.0	609	2 AAY06994	Aay06994 Human alb
43	2068	100.0	609	3 AAB36542	Aab36542 Recombina
44	2068	100.0	609	3 AAY78147	Aay78147 Pre human
45	2068	100.0	609	3 AAB36549	Aab36549 Recombina

ALIGNMENTS

RESULT 1

AAP90389 standard; protein; 388 AA.

XX	AC	AAP90389;
XX	AC	AAP90389;
XX	DT	24-OCT-2003 (revised)
XX	DT	25-MAR-2003 (revised)
XX	DT	01-NOV-1989 (first entry)
XX	DE	N-terminal human serum albumin polypeptide.
XX	DE	N-terminal human serum albumin polypeptide.
XX	OS	Homo sapiens; (Human).
XX	PN	EP322094-A.
XX	PD	28-JUN-1989.
XX	PF	25-OCT-1988; 88EP-00310000.
XX	PR	30-OCT-1987; 87GB-00025529.
XX	PA	(DELZ) DELTA BIOTECHNOLOGY LTD.
XX	PI	Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX	DR	WPI; 1989-186464/26.
XX	PT	New N-terminal fragments of human serum albumin - esp. useful as blood plasma expanders.
XX	PS	Claim 2; Page 9; 20pp; English.
XX	CC	N-terminal fragment of human serum albumin used as plasma expander, or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)
XX	SQ	Sequence 388 AA;

Query Match 100.0%; Score 2068; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 6.4e-192;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DAKHSEVAHRFKDLEENFKALVLIAPQYLCQCFEDHVKLVNEVTFATCVADESAE 60

Db	215	LNEHVCSVIRKFSRNLQATTIIKLSQKLTETANFTIQLALDVALIHRECCQNSLECL	274
QY	255	DDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDMPADLPSLAADFVESKD	314
Db	275	QDGEKWTYICQQNTLSSKIAECCCLPMIQLGFCIIHAENGVKPEGJSLNPSQFLGDRN	334
QY	315	VCKNYAEAKDVFLGMPLYEYARRHPDYSVLLRLAKTYETTLKCCCAADPHECYAKVF	374
Db	335	FAQFSSEKIMFMASFHEYSETHPNLPVSVILRIAKTYQETLEKCSQSGNLPQCQDNLE	394
QY	375	DEFKPLVEEPQNL	387
Db	395	BELOKHIESQAL	407

Search completed: April 19, 2004, 11:59:57
Job time : 151.368 secs

Db 87 DMEGLPQKH-NFSSHCCGKAGFPRLCFFYNKANKVGFLLPPTL---DPEKCKQAKYKNS 142
QY 132 ETLKKLYIYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLPKKDELRLDEGKA 191
Db 143 ESFLHLYMYEVARPNPFVAPVLLAAVAFEEAATTCCEQQOKATCFQAKAAPITQYLKA 202
QY 192 SSQKRLKASLOKGERAFKAWARLSORPFAEFAEVSGLVTLTKVHTECHGDL 251
Db 203 SSSYQRNVGALIKFGPKVLNINAVFSKFPKLGFKDLTLLSDVSSMYEGCGEDV 262
QY 252 ECADRADLAKYICENQDSISSKJKECCPKLLEKSHCIAEVENDEMPADLPSLAADFVE 311
Db 263 HCIRSOQVNVHICSKQDSISSKIVKCEKTLEREACIINANKDDRPEGLSLREAKPTE 322
QY 312 SKDVCNRYAEAKDVLGFLMYIYARHPDYSVLLRLAKTYETTLKCCAAADPHECYA 371
Db 323 SENVOERDSDPKFPAEFIYSRRHPLSTPELLRITKYVMDFLEDCSRENPAAGCYR 382
QY 372 KVFDEF 377
Db 383 HVEDKF 388

RESULT 14
Q8UW06 PRELIMINARY; PRT; 624 AA.
AC Q8UW06; 2003 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DE 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Serum albumin precursor.
OS ALB.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroides; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and Ambystoma texanum";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF2117182; AAL56645.1; --
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 1.
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 624 SERUM ALBUMIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08533BF4953EF7 CRC64;

Query Match 30.5%; Score 630; DB 13; Length 624;
Best Local Similarity 33.7%; Pred. No. 4.1e-46;
Matches 125; Conservative 66; Mismatches 170; Indels 10; Gaps 3;
QY 14 LGENFKALVLIJAFAYLQCCPFEDHVKLVNVEVFATCTVADESAENCDKSLHTLFGDK 73
Db 45 IGVHAKALAMALFQMLSKCPHHEGVQVRVMDIADLCRGAKHGDCGKSVWIIINE 104
QY 74 LCTVATLRETYGEMADCCAKQPERNECFLOHKDNP-NLPLRVPEVDVMTAFHDNEE 132
Db 105 ICTPENPEKYPPHEGCKCKEDPERHKCFIEHKSITDPKERTERYVXKPSPEQICKHAENRD 164
QY 133 TELKKLYIYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLPKDELRLDEGKA 191

Db 165 EPLGHVTHKVASSTHTMYPPAILSLFTLHPDGIIVSHCCCKDEATVGQCLSEKMPAKHEVEH 224
QY 192 SSQKRLKASLOKGERAFKAWARLSORPFAEFAEVSGLVTLTKVHTECHGDL 251
Db 225 VCAVQXHCNVCILQNFNERALRASKAHAHACSKFPHASFENVQRLTDGIVHLHQTCGCGDMM 284
QY 252 ECADRADLAKYICENQDSISSKJKECCPKLLEKSHCIAEVENDEMPADLPSLAADFVE 311
Db 285 ACVAERKLTQTQCEK-----KCCCKPVLSECEIVLPENDEKPADLPSEVRYPD 336
QY 312 SKDVCNRYAEAKDVLGFLMYIYARHPDYSVLLRLAKTYETTLKCCAAADPHECYA 371
Db 337 DPEVCKFFKEEGDAFNGFLCDYAKIHPBSAELNRIASGLEKAYKTCGAGEAHNECIA 396
QY 372 KVFDEFKPLVE 382
Db 397 KBEETLRHEIE 407

RESULT 15
Q8BK65 PRELIMINARY; PRT; 605 AA.
AC Q8BK65; 2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Alpha-fetoprotein precursor.
GN AFP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
DR EMBL; AK076053; BAC36150.1; --
DR MBL; MGI-87951; Afp.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0001542; P:ovulation (sensu Mammalia); IMP.
DR GO; GO:0042448; P:progesterone metabolism; IMP.
DR GO; GO:0019953; P:sexual reproduction; IMP.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 605 AA; 67322 MW; 048B7A4A8B01EA4B CRC64;

Query Match 29.7%; Score 615; DB 11; Length 605;
Best Local Similarity 32.4%; Pred. No. 7.9e-45;
Matches 121; Conservative 80; Mismatches 166; Indels 6; Gaps 3;
QY 16 EENFKALVLIJAFAYLQCCPFEDHVKLVNVEVFATCTVADESAENCDKSLHTLFGDKLC 75
Db 40 EKNVLIATITFTQFVPEATEEE---VNMKTSVDLAAKMKNSGDCLESQSVFLDEIC 95
QY 76 TVATLRETYGEMADCCAKQPERNECFLOHKDNP-NLPLRVPEVDVMTAFHDNEET 134
Db 96 HETELSKYQ-LGSCSGSGVERHQCLLAKKATAPASVFPFQPEAESCAHEBRAV 154
QY 135 LKKLYIYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLPKDELRLDEGKASSA 194
Db 155 MNRFIYVSRRNPFMYAPAILSLAAQYDKVVLACKADNKEECFQTKRASIAKELREGSM 214
QY 195 KQRLKASLOKGERAFKAWARLSORPFAEFAEVSGLVTLTKVHTECHGDLLECA 254


```

RP SEQUENCE FROM N.A.
RA Kim J.G., Noneman D., Vallet J.L., Christenson R.K.:
RT "Mapping of the porcine alpha-fetoprotein (AFP) gene to SSC8.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF517770; AAM66710.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 610 AA; 68624 MW; C985BEAD44963D5E CRC64;

Query Match 33.7%; Score 697.5; DB 6; Length 610;
Best Local Similarity 36.4%; Pred. No. 5.5e-52;
Matches 136; Conservative 74; Mismatches 161; Indels 3; Gaps 3;

QY 16 EENFKALVLIAPFAQLQCCPFEDHVKLVNEVTEFAKTCVADSESAENCDKSLHTLFGDKLC 75
DB 40 EENLVDLATIFPAQFVHGATYTEIRKANDVLAVIKKPTDGHPTGCLENQMSAFLEBIC 99
QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDDNP-NLPLRVLRPEVDVMTAFHDNEETP 134
DB 100 HEEIEPEKYG-LSHCCSQSGEERHNCFLARKKAAPASIPFFQVPEFVTSKAYEENRELF 158
QY 135 LKLYEIAARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAAACLLPKLDELDRDEGKASSA 194
DB 159 MTEYIIEIARRHPFLYAPILSLAAGYDKMIIPCCKAENAVECFQTKAASITKELRESSL 218
QY 195 KQRLKASLOKQGERAFKAWAVARLSORFPKAFPAEVSKLVTDLTKVHTCCCHGDLLECA 254
DB 219 LNCHMCTVMRQFGARTFRAITVTKLSQKPKANFTIQLKLVLDVAHIHECCRCGNVLESL 278
QY 255 DDRADLAKYICENODSISKKLKECCPKLLEKSHCIAEVDENDEMPADLPSLAADFVBSKV 313
DB 279 QDAERVVSVCSQDITLSSKIAECCKLTLELGQCIIHAENDDKPEGLSPNLNLFGER 338
QY 314 DVCKNYAEAKDVLGMFLYEAARRHPDYVSVLLLRLLAKTYETTLKCCAAADPHECYAKV 373
DB 339 DFNQLSREKDLGMAFTVEYSRRHPKLAVPVLIRVAKGYQELLEKCSQSENPLECQDKG 398
QY 374 FDEFKPLVEEPQNL 387
DB 399 EEELEKVIQESQAL 412

RESULT 12
Q7TSF3 PRELIMINARY; PRT; 608 AA.
AC Q7TSF3; 2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DR GO; GO:0005615; C:extracellular space; IEA.
DE Alpha-fetoprotein.
OS Marmota monax (Woodchuck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Marmota.
OX NCBI_TaxID=9995;
FT NON_TER 1
SQ SEQUENCE 608 AA; 67962 MW; EB46D7A37382B222 CRC64;

Query Match 33.2%; Score 686.5; DB 11; Length 608;
Best Local Similarity 36.8%; Pred. No. 5e-51;

QY 16 EENFKALVLIAPFAQLQCCPFEDHVKLVNEVTEFAKTCVADSESAENCDKSLHTLFGDKLC 75
DB 27 DENTAYLAIAPFSQVQEAASFDEVELVKWMLDRDRCWADNTLPESKSTANDAIQDWLC 86
QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDDN-----PNLPLRVLRPEVDVMTAFHDNE 131

Matches 136; Conservative 66; Mismatches 167; Indels 1; Gaps 1;

QY 16 EENFKALVLIAPFAQLQCCPFEDHVKLVNEVTEFAKTCVADSESAENCDKSLHTLFGDKLC 75
DB 40 EENLVDLATIFPAQFVHGATYTEIRKANDVLAVIKKPTDGHPTGCLENQMSAFLEBIC 99
QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDDNP-NLPLRVLRPEVDVMTAFHDNEETP 135
DB 100 HEEIEPEKYG-LSDCCSRGSEERHSCLPAPHKAAAPALPFPQAPPEPATSKAYEEDRET 158
QY 136 KLYEIAARRHPFYFAPPELLFFAKRYKAAFTCCQAADKAAACLLPKLDELDRDEGKASSA 195
DB 159 NKCIYIISRRHPFLYAPILSLAAGYDKMIIPCCKAENAVECFQTKAASITKELRESSL 218
QY 196 QRLKASLOKQGERAFKAWAVARLSORFPKAFPAEVSKLVTDLTKVHTCCCHGDLLECAD 255
DB 219 NQHVCAVKTETFEPLIFQAITVAKLSQKFPFRANFTETQKLVLDVAHIHECCRCGNVLECLQ 278
QY 256 DRADLAKYICENODSISKKLKECCPKLLEKSHCIAEVDENDEMPADLPSLAADFVBSKV 315
DB 279 DGEKINSYICSRGALSSKIAECCKLTLELGQCIIHAENDDKPEGLSPTLNRFGLDRDP 338
QY 316 CKNYAEAKDVLGMFLYEAARRHPDYVSVLLLRLLAKTYETTLKCCAAADPHECYAKV 375
DB 339 NQSSAERNIFLASFVHYSRRHPKLAVPVLIRVAKGYQELLEKCAQSEKPLECQDKG 398
QY 376 EFKPLVEEPQ 385
DB 399 ELQKTYQESQ 408

RESULT 13
Q8R0J9 PRELIMINARY; PRT; 417 AA.
AC Q8R0J9; 2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DR GO; GO:0005615; C:carrier activity; IEA.
DE Hypothetical protein (Fragment).
GN AFM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
FT [1]
RP SEQUENCE FROM N.A.
RC Tissue-Kidney;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026681; AAN26681.1; -.
DR MGD; MGI:2429409; Afm.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 417 AA; 47667 MW; AF8F4FF0A76A92B3 CRC64;

Query Match 32.7%; Score 677; DB 11; Length 417;
Best Local Similarity 37.4%; Pred. No. 2.1e-50;
Matches 137; Conservative 64; Mismatches 157; Indels 8; Gaps 3;

QY 16 EENFKALVLIAPFAQLQCCPFEDHVKLVNEVTEFAKTCVADSESAENCDKSLHTLFGDKLC 75
DB 27 DENTAYLAIAPFSQVQEAASFDEVELVKWMLDRDRCWADNTLPESKSTANDAIQDWLC 86
QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDDN-----PNLPLRVLRPEVDVMTAFHDNE 131

```

```

QY 116 VRPEVDVMTAFHDNEETFLKKYLYEARRHPYFYAPPELLFFAKRYKAAFTCCCAADKA 175
Db 148 VRPAEQIQCKDAENRGPILLARYIFMLAIGHPMYIPAILGFAQFDGIVSHCKDVETA 207
QY 176 A-CLLPKLDLRDEKASAKQRLKASLOKFGERAFAKAWARLRSQRPFAEYVSKL 234
Db 208 GQCFNDKMPKHEQVEYVYCALQKNCYILQDFKERALTAYKAVQASQKFFPLASFENQII 267
QY 235 VTDLKVTECHGCHLLECCADRADLAKVICHENODSISKKKECKEPLLEKSHCIAEVE 294
Db 268 VPDTHLHQTCGCGMMACHLERMKLTAKICEKDELATHLKECCDKELLERSACIIRLP 327
QY 295 NDEMPADIPSLLAADFVSEKOVCKVYAEAKDVLGMLFYEARHPDYVSWILLRLAKTYE 354
Db 328 NDKPADLSPKPVHYIDDPEVCKLYTEGGDTFMGRFLYECARRHODYSPEMLLRWGSGYE 387
QY 355 TTEKCCAAADPHECYAKVDFEFKLVPEPQNL 387
Db 388 BFLKCCCAAGHNECLAETESLKEIEBSVTL 420

RESULT 9
Q8CG74 PRELIMINARY; PRT; 205 AA.
AC Q8CG74; (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Albumin (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvevTacrBr;
RA Van Reeth T., Dreze P.L., Gabant P., Szpirer C., Szpirer J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; A2277794; CAC81903.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 1.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 1.
FT NON_TER 205
SQ SEQUENCE 205 AA; 23624 MW; 0A982E85C7474BF9 CRC64;

Query Match 35.7%; Score 739; DB 11; Length 205;
Best Local Similarity 69.6%; Pred. No. 3.7e-56;
Matches 126; Conservative 26; Mismatches 29; Indels 0; Gaps 0;

QY 1 DAHKEVNRFDLGEENFKALVLIAPYLOQCPEDHVKLINVEVTEFAKTCVADESAG 60
Db 25 EAHKSEIARHYNDLGEQHFQGLVLIAPYLOQCSYDEHAKLVQEVTEFAKTCVADESAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECEFLQHKDNPMLPRVREV 120
Db 85 NCDKSLHTLFGDKLCAIPNLRENYGELADCCCTQEPERNECEFLQHKDNPSPFPEREA 144
QY 121 DVNCTAFHDNEETFLKKYLYEARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACILP 180
Db 145 EAMCTSFKENPTTFMGHYLHEVARRHPYFYAPPELLYAEQYNEILTQCCAEADKESCLTP 204
QY 181 K 181
Db 205 K 205

RESULT 10
Q8MJU5 PRELIMINARY; PRT; 609 AA.
ID Q8MJU5

```

```

AC Q8MJU5;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Alpha-fetoprotein.
GN AFP.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Furuichi M., Neo S., Hisasue M., Teuchiya R., Watanabe M.,
RA Hashizaki K., Hisamatsu S., Yamada T.;
RT "Canine alpha-fetoprotein cDNA."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB089789; BAC07513.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 609 AA; 68782 MW; BE4B8250C5AF2AF0 CRC64;

Query Match 34.4%; Score 712; DB 6; Length 609;
Best Local Similarity 37.3%; Pred. No. 3.1e-53;
Matches 139; Conservative 70; Mismatches 162; Indels 2; Gaps 2;

QY 16 BENFKALVLIAPYLOQCPEDHVKLINVEVTEFAKTCVADESAGNCDKSLHTLFGDKLC 75
Db 40 EMNLVDLATIFFAQFVQEATYKEVSKMKDKILTVIERKSTGSEQFGGLENQLFAFLEEIC 99
QY 76 TVATLRETYGEMADCCAKQEPERNECEFLQHKD-DNPNLPRLVPEVDVMTAFHDNEET 134
Db 100 HEKEISEKYG-LADCCSQREERHNCFLAHKKAAPPSIPFPQVAEPVTSCKAYEENRDMF 158
QY 135 LKKYLYEARRHPYFYAPPELLFFAKRYKAAFTCCCAADKAACILPDLRLDEGKASSA 194
Db 159 MNRYIYEIARRHPFLYAPTILSLAAHYGKIIP:CCKAENAVECFQTKTSLITKELESSL 218
QY 195 KQRLKCAKSLQFGERAFKAWARLRSQRPFAEYVSKLVTLTKVHTSCCHGLLECA 254
Db 219 LNQHICAVMENFGPRTFRAITVTKLSQFSKANFTIQKLVDVAIHBECCRGVLECL 278
QY 255 DADRADLAKVICHENODSISKKKECKEPLLEKSHCIAEVENDEMPADLPSLAADFVSKD 314
Db 279 QDGEKIMSYICSQODILSSKIADCCCKLPILLEGQCIHAENDGKPEGLSPNLNRFLEED 338
QY 315 VCKNYAEAKDVLGMLFYEARHPDYVSVLLRLAKVTEVTTLEKCCAAADPHECYAKVF 374
Db 339 FNQSSREKDLFWARFYEYRRRTKLVAPVVLVAVKGYOELLEKCSQSENPLECQDKGE 398
QY 375 DEFKLVPEPQNL 387
Db 399 EELEKYIQESQAL 411

RESULT 11
Q8MJU6 PRELIMINARY; PRT; 610 AA.
ID Q8MJU6
AC Q8MJU6; (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Alpha-fetoprotein.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;

```

Pfam: PF00273; transport_prot; 2.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR SMART; SM00103; ALBUMIN; 2.
 DR PROSITE; PS00212; ALBUMIN; 2.
 SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 45.3%; Score 936; DB 4; Length 396;
 Best Local Similarity 49.9%; Pred. No. 6.7e-73;
 Matches 196; Conservative 42; Mismatches 121; Indels 34; Gaps 7;

QY 1 DAKSEVAHRFKDLGEENFKALVLIAPAYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 25 DAKSEVAHRFKDLGEENFKALVLIAPAYLOQCFFEDHVKLVNEVTEFAKTCVADESAAE 84

QY 61 NCKSKLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOKHNDPNLRLVRPEV 120
 DB 85 NCKSKLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOKHNDPNLRLVRPEV 144

QY 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFAPPELLFFAKRYKAAATECCOAAADKAAACLLP 180
 DB 145 DVNCTAFHDNEETFLKKYLIEIARRHPYFAPPELLFFAKRYKAAATECCOAAADKAAACLLP 183

QY 181 KLDELRLD--EGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAVSKLVTDL 238
 DB 184 VFDEFKPLVEEPONLIQ--NCLFQGLGEYFQNALVRYTKVPQSTPTLVEVSRL 241

QY 239 TKVHTBCC---HGDILLECADDRADLA-KYIC--ENQDISSKLKECCCKEPLLEKSHCIAE 292
 DB 242 GKVSCKCKHPEAKRPECAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESLVNRRPQPSA 301

QY 293 VENDEMPADLPSLAADFVSKDVKYAKVFLGMFLYIYARRHPDYVSVLLRLAKT 352
 DB 302 LEVDETYVPEKFAETFTFHADICTLSEKQRIKQTALVELVHKHKAKEQLKAVMDD 361

QY 353 YETTLKCCAAADPHECYAKVDFEFKPLVEEPQ 385
 DB 362 FAAFVEKCKCKADDKTCFA---EKGKLVAAEQ 391

RESULT 7
 Q8JIA9 PRELIMINARY; PRT; 527 AA.
 ID Q8JIA9
 AC Q8JIA9;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Serum albumin (Fragment).
 OS Sphenodon punctatus (Hatteria) (Tuatara).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
 OC NCBI_TaxID=8508;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;
 RT "Partial coding sequence for Sphenodon punctatus 68 kDa albumin."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF375971; AAM46104.1;
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; P:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 FT NON_TER 1
 SQ SEQUENCE 527 AA; 59711 MW; C62B799B387F5929 CRC64;

Query Match 38.2%; Score 791; DB 13; Length 527;
 Best Local Similarity 45.3%; Pred. No. 3.6e-60;
 Matches 148; Conservative 53; Mismatches 124; Indels 2; Gaps 2;

QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOKHNDPNLRLVRPEV 120
 DB 5 CLASLDITFLDEICHBEFGAAY-DLAACCAAEVERKECLLAHKNATGFIAPFORPGI 63

QY 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFAPPELLFFAKRYKAAATECCOAAADKAAACLLP 180
 DB 64 EVSKLYQDDRLTLGNLYIYARRHPYLOVPVPATASLYDEALTKCCTQADKATCFHP 123

QY 181 KLDELRLDDEGKASSAKORLKASLOKFGERAFAKAWAVARLSORPPKAEFAVSKLVTDLTK 240
 DB 124 RIPLLEIYKMSNGIQENTCGILKKGERTLKATKLVMQSKPKADFATINKLVEDITH 183

QY 241 VHTCCGDLLECADRADLAKYICENQDISSKLKECCCKEPLLEKSHCIAEAVENDEMPA 300
 DB 184 MTECCRGDTLECLRDEALTEYTCSHKOAISSKLPTCCCKSVLEGECLVREDDDKPA 243

QY 301 DLPSLAADFVSKDVKYAKVFLGMFLYIYARRHPDYVSVLLRLAKTYETTLK 360
 DB 244 DLSERIAEYIEDPHVDHLAKEQDAFLAKFLYYSRRHPSELSTQILLGVGKGYQELLERC 303

QY 361 CAAADPHECYAKVDFEFKPLVEEPQNL 387
 DB 304 CKTDNPPECYGOAEADLKGHIAQFQEL 330

RESULT 8
 Q8UW05 PRELIMINARY; PRT; 626 AA.
 ID Q8UW05
 AC Q8UW05;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Ambystoma maculatum (Spotted salamander).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
 OC Ambystoma.
 OC NCBI_TaxID=43114;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Haverfield E.V., Uzzell T., Spolsky C.M., Bazartseren B.;
 RT "Serum albumin of the mole salamanders Ambystoma maculatum and Ambystoma texanum."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF217183; AAL56646.1;
 DR GO; GO:0005615; C:extracellular space; IEA.
 DR GO; GO:0005386; P:carrier activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 2.
 KW SIGNAL.
 FT CHAIN 1 24 POTENTIAL.
 FT SIGNAL 25 626 SERUM ALBUMIN.
 SQ SEQUENCE 626 AA; 70677 MW; 9D66F57F174AC23F CRC64;

Query Match 36.9%; Score 764; DB 13; Length 626;
 Best Local Similarity 38.9%; Pred. No. 9.7e-58;
 Matches 153; Conservative 64; Mismatches 170; Indels 6; Gaps 3;

QY 1 DAKSEVAHRFKDL----GEENFKALVLIAPAYLOQCFFEDHVKLVNEVTEFAKTCVAD 56
 DB 28 EGHVDNPPHLIGDLIPMIGVDNKGSLVLAASQMLPLCPYEHLQVRVEDVNAQIADLCAG 87

QY 57 ESEAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOKHNDPNLRL 115
 DB 88 ARHANCAPSPWITLDELCKKPENAEKYFFHQECCCKEDPERHKCFVEHKANHELIKY 147

```
QY 361 CAADPHCYAKVDFEKPVEEPQNL 387
DB 385 CEAENPRACYGVLAFAEPQVEEPQNL 411

RESULT 4
Q8C7C7 PRELIMINARY; PRT; 576 AA.
AC Q8C7C7;
DT 01-WAR-2003 (TREMELrel. 23, Created)
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22334683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK050644; BAC34360.1; -.
DR MGD; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
DR NON_TER 1
FT SEQUENCE 576 AA; 65002 MW; F85733B99AE37F04 CRC64;
SQ

Query Match 74.6%; Score 1542; DB 11; Length 576;
Best Local Similarity 71.5%; Pred. No. 2.8e-125;
Matches 271; Conservative 55; Mismatches 53; Indels 0; Gaps 0;

QY 9 HRFKDLGEENKALVIAFAQVLCQCPREDHVKLAVNEVTERAKTCVADESACNCKSLHT 68
DB 1 NRYNDLGEQHFKGLVIAFSQVLOKCSYDEHAKLVQEVTDFAKTCVADESACNCKSLHT 60

QY 69 LFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPRLVPEVDMCTAFH 128
DB 61 LFGDKLCAIPNLFRENYGELADCTKQEPERNECFLOHKDDNPSLPFPERPEAEAMCTGFK 120

QY 129 DNEETFLKYLVEIARHRYFYAPELLFFAKRYKAAPTECCQAADKAACLLPKLDELDE 188
DB 121 ENPTFMGHLVEHARRHYFYAPELLFYAEQNEILTOCCAEADKESCLTLPKLDGVKEK 180

QY 189 GRASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAVSKLVITDLTKVHTCCCHG 248
DB 181 ALVSSVRQRMKSSCMQKFGERAFKAWAVARLSQTFPNADFARITKLTATLTKVKECCCHG 240

QY 249 DLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADLPSLAAD 308
DB 241 DLLECADRADLAKYICENQATISSKLTQCCDKPLLKKAHCLSEVEHDTMPADLPAIAD 300

QY 309 FVSKDKVCKNYAAKQVFLGMFLYEVARRHPDYSVVLRLAKYVTTLEKCCAAADPHE 368
DB 301 FVEDQEVCKNYAAKQVFLGTFLYEVSRHPDYSVSLRLAKYATLEKCCAAENPPA 360

QY 369 CYGTVLAEFQPLVEEPQNL 387
DB 361 CYGTVLAEFQPLVEEPQNL 379
```

RESULT 5

```
Q86YGO PRELIMINARY; PRT; 417 AA.
ID Q86YGO;
AC Q86YGO;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
DR SEQUENCE 417 AA; 47360 MW; 16E764833EEF4E8D CRC64;
SQ

Query Match 46.2%; Score 956; DB 4; Length 417;
Best Local Similarity 90.6%; Pred. No. 1.3e-74;
Matches 183; Conservative 3; Mismatches 10; Indels 6; Gaps 1;

QY 186 RDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRFPAEFAVSKLVITDLTKVHTC 245
DB 24 RDAHSEVAHR-----FKDLGEENKALVIAFAQVLCQCPREDHVKLAVNEVTERAKTCVADESACNCKSLHT 60

QY 246 CHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADLPSL 305
DB 78 CHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADLPSL 137

QY 306 AADFVESKDVCKNYAAKQVFLGMFLYEVARRHPDYSVVLRLAKYVTTLEKCCAAAD 365
DB 138 AADFVESKDVCKNYAAKQVFLGMFLYEVARRHPDYSVVLRLAKYVTTLEKCCAAAD 197

QY 366 PHECYAKVDFEKPVEEPQNL 387
DB 198 PHECYAKVDFEKPVEEPQNL 219

RESULT 6
Q81UK7 PRELIMINARY; PRT; 396 AA.
ID Q81UK7;
AC Q81UK7;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR SEQUENCE 396 AA; 47360 MW; 16E764833EEF4E8D CRC64;
SQ
```

```
Db 121 DAMCTAFHENEQRFGLGKYLIEIARRHPYFYAPPELLYAAEYRGVFTCECEAAADKAACTLP 180
Qy 181 KLDELDRSGKASSAKORLKASLOKFGGERAFKAWAVARLSQRFPPKABFAEVSCLVTLTK 240
Db 181 KYDALREKVLASSAKERLKASLOKFGGERAFKAWAVARLSQRFPPKABFAEVSCLVTLTK 240
Qy 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLECCCKPILLESKHCIAEVENDEMPA 300
Db 241 IHKECCHGDLLECCADDDRADLAKYICENQDSISSKLECCCKPILLESKHCIAEVENDEMPA 300
Qy 301 DLPSLAADPVEKDVCKYKNAEAKDVLGMLFYEARHPDYVSULLLRLAKYETTLTK 360
Db 301 DLPLLAADPVEKDVCKYKNAEAKDVLGMLFYEARHPDYVSULLLRLAKYETTLTK 360
Qy 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
Db 361 CATDPPACVYAHVDFEFKPLVEEPHNL 387
```

RESULT 2

```
Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
SEQUENCE FROM N.A.
RA Osman A., Asahi H., Staderker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RT component in murine infection with Schistosoma mansoni.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418550; AAL08579.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EABB28E1C66B54 CRC64;
```

```
Query Match 81.3%; Score 1681; DB 5; Length 608;
Best Local Similarity 76.5%; Pred. No. 2.6e-137;
Matches 296; Conservative 53; Mismatches 38; Indels 0; Gaps 0;
Qy 1 DAHSEVAHRFKDLGRENFKALVLIAPAOYLQOCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHSEIAHRYNDLGEQHFGLVLIAPSOPLQCPYEEHVKLVEVTDFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 85 NCDKSLHTLFGDKLCAIPLTLDSDYSELADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 144
Qy 121 DVMTAFHDNEETFLKYLIEIARRHPYFYAPPELLYFAKRYKAAFTCCQADKAACTLP 180
Db 145 EAMCTSFQENAVTFMGHYLHEVARHPYFYAPPELLYAEKYAISIMTECCGEADKAACTIP 204
Qy 181 KLDELREGKASSAKORLKASLOKFGGERAFKAWAVARLSQRFPPKABFAEVSCLVTLTK 240
Db 205 KLDALKEKALASSVQRLKSSQRFQORAFKAWAVARLSQRFPPKADFAEITKLTDLTK 264
Qy 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLECCCKPILLESKHCIAEVENDEMPA 300
Db 265 LTERCCHGDLLECCADDDRADLAKYICENQDSISSKLECCCKPILLESKHCIAEVENDEMPA 324
```

```
Qy 301 DLPSLAADPVEKDVCKYKNAEAKDVLGMLFYEARHPDYVSULLLRLAKYETTLTK 360
Db 325 DLPSLAADPVEKDVCKYKNAEAKDVLGMLFYEARHPDYVSULLLRLAKYETTLTK 384
Qy 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
Db 385 CABADSPSACYKVLDFEFQPLVEEPKNL 411
RESULT 3
Q8C7H3 PRELIMINARY; PRT; 608 AA.
AC Q8C7H3
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Albumin 1.
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RC MEDLINE=23354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGD; MGI-87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292F600EED3A61B4 CRC64;
```

```
Query Match 76.6%; Score 1584; DB 11; Length 608;
Best Local Similarity 71.8%; Pred. No. 6.8e-129;
Matches 278; Conservative 56; Mismatches 53; Indels 0; Gaps 0;
Qy 1 DAHSEVAHRFKDLGRENFKALVLIAPAOYLQOCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHSEIAHRYNDLGEQHFGLVLIAPSOPLQCPYEEHVKLVEVTDFAKTCVADESAAE 84
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
Db 85 NCDKSLHTLFGDKLCAIPLTLDSDYSELADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 144
Qy 121 DVMTAFHDNEETFLKYLIEIARRHPYFYAPPELLYFAKRYKAAFTCCQADKAACTLP 180
Db 145 EAMCTSFQENAVTFMGHYLHEVARHPYFYAPPELLYAEKYAISIMTECCGEADKAACTIP 204
Qy 181 KLDELREGKASSAKORLKASLOKFGGERAFKAWAVARLSQRFPPKABFAEVSCLVTLTK 240
Db 205 KLDGVEKALVSSVQRKSSQRFQORAFKAWAVARLSQRFPPKADFAEITKLTDLTK 264
Qy 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLECCCKPILLESKHCIAEVENDEMPA 300
Db 265 VNKECCHGDLLECCADDDRADLAKYICENQATISSKLOCTCCDKPLKKAHCISEVEDTMPA 324
Qy 301 DLPSLAADPVEKDVCKYKNAEAKDVLGMLFYEARHPDYVSULLLRLAKYETTLTK 360
Db 325 DLPSLAADPVEKDVCKYKNAEAKDVLGMLFYEARHPDYVSULLLRLAKYETTLTK 384
```

Q9YGH6 rana shqipe
Q91134 naja naja
Q8JIA7 sphendonon p
Q8JIA8 hoplodactyl
Q8W615 gallus gall
Q7T897 mus musculu
Q42279 petromyzon
Q95MC2 equus caball
Q8J2S7 helicobacte
Q74669 pneumocysti
Q9JMA8 helicobacte
Q9TRAS oryctolagus
Q9ZLV0 helicobacte
Q25262 helicobacte
Q74668 pneumocysti
Q8YK55 anabaena sp
Q9U540 drosophila
Q9MS79 drosophila
Q9VU94 drosophila
Q9UVY2 pneumocysti
Q9QVA1 rattus sp.
Q01794 pneumocysti
Q9V688 drosophila
Q01828 pneumocysti
P78734 pneumocysti
Q12075 pneumocysti
Q8ID09 plasmodium
Q61194 mus musculu
Q61182 mus musculu

17 591 28.6 603 13 Q9YGH6
18 572.5 27.7 614 13 Q91134
19 562 27.2 400 13 Q8JIA7
20 409.5 19.8 406 13 Q8JIA8
21 352.5 17.0 484 13 Q9W615
22 341 16.5 491 11 Q7T897
23 275 13.3 551 13 Q42279
24 162.5 7.9 44 6 Q95MC2
25 159.5 7.7 807 2 Q8J2S7
26 152 7.4 1025 2 Q74669
27 150 7.3 1723 2 Q9JMA8
28 149 7.2 40 6 Q9TRAS
29 149 7.2 1819 16 Q9ZLV0
30 148.5 7.2 1927 16 Q25262
31 146.5 7.1 1028 3 Q74668
32 121.5 5.9 1348 16 Q8YK55
33 121.5 5.9 3843 5 Q9U540
34 120.5 5.8 661 5 Q9MS79
35 120.5 5.8 3843 5 Q9VU94
36 114 5.5 1092 3 Q9UVY2
37 111 5.4 21 11 Q9QVA1
38 111 5.4 1005 3 Q01794
39 110 5.3 680 5 Q9V688
40 110 5.3 1065 3 Q01828
41 109.5 5.3 1070 3 P78734
42 109 5.3 1083 3 Q12075
43 109 5.3 3954 5 Q8ID09
44 108.5 5.2 1509 11 Q61194
45 108.5 5.2 1658 11 Q61182

ALIGNMENTS

RESULT 1

Q7YSG3 PRELIMINARY; PRT; 584 AA.
AC Q7YSG3
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Albumin (Fragment)
GN ALB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
RX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Reininger R., Swoboda I., Bohle B., Hauswith A.W., Valent P.,
RA Rumpold H., Valenta R., Spitzauer S.;
RT "Escherichia coli expression and purification of recombinant cat
RT albumin: IGE recognition, induction of basophil activation and
RT lymphoproliferative responses in atopic patients";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ487677; CAD32275.1; -
FT NON_TER 1
SQ SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 83.2%; Score 1721; DB 6; Length 584;
Best Local Similarity 80.1%; Pred. No. 8.4e-141;
Matches 310; Conservative 36; Mismatches 41; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDLGEENFKALVLIARFAYLQQCPFFEDHVKLVNEVTEFAKTCVADESA 60
Db 1. EAHQSEIAHRNDLGESEFRGLVLFVAFSOLQQCPFFEDHVKLVNEVTEFANGCVAQSA 60
QY 61 NCKSLHTLFGDKLCTVATLTRETYGEMADCCAKPEPERNECFLOHKDNPPLVREV 120
Db 61 NCKSLHTLFGDKLCTVATLTRETYGEMADCCAKPEPERNECFLOHKDNPPLVREV 120
QY 121 DVMCTAFDNEETFLKYLVEIARHPFYFAPELLFFAKRYKAATFECQADKACLLP 180

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 149.368 Seconds

(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18_COPY_1_387

Perfect score: 2068

Sequence: 1 DAHSEVAHRFKDLGEENFK.....ECYAKVFDFEKLVEEPQNL 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mbc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1721	83.2	584	5 Q7YSG3	Q7YSG3 felis silve
2	1681	81.3	608	5 Q95VB7	Q95VB7 schistosoma
3	1584	76.6	608	11 Q8C7H3	Q8C7H3 mus musculu
4	1542	74.6	576	11 Q8C7C7	Q8C7C7 mus musculu
5	956	46.2	417	4 Q86YGO	Q86YGO homo sapien
6	936	45.3	396	4 Q8IUK7	Q8IUK7 homo sapien
7	791	38.2	527	13 Q8JIA9	Q8JIA9 sphendonon p
8	764	36.9	626	13 Q8UW05	Q8UW05 ambystoma m
9	739	35.7	205	11 Q8CG74	Q8CG74 mus musculu
10	712	34.4	609	6 Q8MUU5	Q8MUU5 canis famli
11	697.5	33.7	610	6 Q8MJ76	Q8MJ76 sus scrofa
12	686.5	33.2	608	11 Q7TSF3	Q7TSF3 marmota mon
13	677	32.7	417	11 Q8R0J9	Q8R0J9 mus musculu
14	630	30.5	624	13 Q8UW06	Q8UW06 ambystoma t
15	615	29.7	605	11 Q8BK65	Q8BK65 mus musculu
16	608	29.4	605	11 Q8BK56	Q8BK56 mus musculu

Qy	241	VHTECCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA	300
Db	241	VHTECCHGDLLECCADDRADLAKYICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPA	300
Qy	301	DLPSLAADFVESKDCVCKNVAEAKDVLGMLFVEYARRHPDYSVLLLRLLAKTYETTTLEKC	360
Db	301	DLPSLAADFVESKDCVCKNVAEAKDVLGMLFVEYARRHPDYSVLLLRLLAKTYETTTLEKC	360
Qy	361	CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKKVPQVST	420
Db	361	CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNQALLVRYTKKVPQVST	420
Qy	421	PTLVEVSRNLGKVGSKCKKHPKAKMPCAEDVLSVVLNQLCVLHEKTPVSDRVTCKCTES	480
Db	421	PTLVEVSRNLGKVGSKCKKHPKAKMPCAEDVLSVVLNQLCVLHEKTPVSDRVTCKCTES	480
Qy	481	LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTALVELVKKPKAT	540
Db	481	LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKERQIKKQTALVELVKKPKAT	540
Qy	541	KEQLKAVMDDFAAFVEKCKKADDKETCPAEKGKLVAAASQAALGL	595
Db	541	KEQLKAVMDDFAAFVEKCKKADDKETCPAEKGKLVAAASQAALGL	595

Search completed: April 19, 2004, 12:54:56
 Job time : 247.316 secs

Db 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLRLAKTYETILEKC 360
Qy 361 CAAADPHCYAKVDFEFKPLVBPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
Db 361 CAAADPHCYAKVDFEFKPLVBPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Qy 481 LVNRRPCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Qy 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 14

US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-319-263-1

Query Match 100.0%; Score 3103; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEBTEFLKKYLYEIARRHPYFYAPELLFFAKRYKAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEBTEFLKKYLYEIARRHPYFYAPELLFFAKRYKAFTCCQAADKAAACLLP 180
Qy 181 KLDELDRDEGKASSAKQRLKCAASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
Db 181 KLDELDRDEGKASSAKQRLKCAASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCECKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCECKPLLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLRLAKTYETILEKC 360

Db 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLYIYARRHPDYSVLLRLAKTYETILEKC 360
Qy 361 CAAADPHCYAKVDFEFKPLVBPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
Db 361 CAAADPHCYAKVDFEFKPLVBPQNLIKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCKCTES 480
Qy 481 LVNRRPCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETVYPKEFNAETFTFHADICTLSEKERQIKKQTALVELVGHKPKAT 540
Qy 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 15

US-10-319-263-2
; Sequence 2, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Lau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MOD_RES
; LOCATION: (1)..(585)
; OTHER INFORMATION: ACETYLYATION
US-10-319-263-2

Query Match 100.0%; Score 3103; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRPKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEBTEFLKKYLYEIARRHPYFYAPELLFFAKRYKAFTCCQAADKAAACLLP 180
Db 121 DVMCTAFHDNEBTEFLKKYLYEIARRHPYFYAPELLFFAKRYKAFTCCQAADKAAACLLP 180
Qy 181 KLDELDRDEGKASSAKQRLKCAASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
Db 181 KLDELDRDEGKASSAKQRLKCAASLOKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240
Db 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240
Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Qy 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNQALLVRYTKKVPQVST 420
Db 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNQALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHETPVS DRVTKCCTES 480
Db 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHETPVS DRVTKCCTES 480
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Qy 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEAGKGLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEAGKGLVAASQAALGL 585

RESULT 12
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match 100.0%; Score 3103; DB 13; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGGENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGGENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240
Db 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240
Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360

Db 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240
Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Qy 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNQALLVRYTKKVPQVST 420
Db 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKONCELFEQLGEYKFNQALLVRYTKKVPQVST 420
Qy 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHETPVS DRVTKCCTES 480
Db 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVNLQCVLHETPVS DRVTKCCTES 480
Qy 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVGHKPKAT 540
Qy 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEAGKGLVAASQAALGL 585
Db 541 KEOLKAVMDDFAAFVEKCKKADDDKTCFAEAGKGLVAASQAALGL 585

RESULT 13
US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5

Query Match 100.0%; Score 3103; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGGENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGGENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240
Db 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLTk 240
Qy 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLFYEARHPDYSVLLRLAKTYETTTLEKC 360

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEYIARRHPYFYAPELLFFAKRYKAAFTTECCQAAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEYIARRHPYFYAPELLFFAKRYKAAFTTECCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTTLEKC 360
QY 361 CAAADPHECVAKVDFEFPKPLVEBPQNLKQNCLEFQOLGEYKFPQNALLVRYTKKVPQVST 420
DB 361 CAAADPHECVAKVDFEFPKPLVEBPQNLKQNCLEFQOLGEYKFPQNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALVDETVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
DB 481 LVNRRPCFSALVDETVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
QY 541 KEQLKAVMDDPFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585
DB 541 KEQLKAVMDDPFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585
RESULT 10
US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11
Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEYIARRHPYFYAPELLFFAKRYKAAFTTECCQAAADKAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEYIARRHPYFYAPELLFFAKRYKAAFTTECCQAAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLQKGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTTLEKC 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTTLEKC 360
QY 361 CAAADPHECVAKVDFEFPKPLVEBPQNLKQNCLEFQOLGEYKFPQNALLVRYTKKVPQVST 420
DB 361 CAAADPHECVAKVDFEFPKPLVEBPQNLKQNCLEFQOLGEYKFPQNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALVDETVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
DB 481 LVNRRPCFSALVDETVVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVGHKPKAT 540
QY 541 KEQLKAVMDDPFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585
DB 541 KEQLKAVMDDPFAAFVEKCKCKADDKETCFABEGKLVAAASQAALGL 585
RESULT 11
US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31
Query Match 100.0%; Score 3103; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYLQCCPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

```
; ORGANISM: Homo Sapiens
US-09-832-501-18

Query Match      100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
Qy 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAAADKAAACLLP 180
Db 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAAADKAAACLLP 180
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLEKC 360
Qy 361 CAAADPHECVAKYDFEFKPLVEBPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
Db 361 CAAADPHECVAKYDFEFKPLVEBPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKQTALVELVKGKPKAT 540
Db 481 LVNRRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKQTALVELVKGKPKAT 540
Qy 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 8
US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-118-18

Query Match      100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
Qy 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAAADKAAACLLP 180
Db 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAAADKAAACLLP 180
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLEKC 360
Qy 361 CAAADPHECVAKYDFEFKPLVEBPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
Db 361 CAAADPHECVAKYDFEFKPLVEBPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKQTALVELVKGKPKAT 540
Db 481 LVNRRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKQTALVELVKGKPKAT 540
Qy 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585

RESULT 9
US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match      100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDLGSENFKALVLIAPAYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVREPV 120
Qy 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAAADKAAACLLP 180
Db 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCCAAADKAAACLLP 180
Qy 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Db 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSORPPKAEFAEVSCLVTDLT 240
Qy 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLEKC 360
Qy 361 CAAADPHECVAKYDFEFKPLVEBPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
Db 361 CAAADPHECVAKYDFEFKPLVEBPQNLIKONCELFEQLGEYKFQNALLVRYTKVPQVST 420
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
Qy 481 LVNRRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKQTALVELVKGKPKAT 540
Db 481 LVNRRPCFSALEVEDTVVPKEFNAETTFHADICTLSEKERQIKQTALVELVKGKPKAT 540
Qy 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
Db 541 KEQLKAVMDPFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
```

```
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-117-18.

Query Match      100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHSEVAHFRKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFDGKLCITVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFDGKLCITVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNBEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHDNBEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPILLESKHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPILLESKHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360
DB 301 DLPSLAADPVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVFDEPKPLVEEPQNLIKONCELFEQLGKGFQNALLVRYTKVPQVST 420
DB 361 CAADAPHECYAKVFDEPKPLVEEPQNLIKONCELFEQLGKGFQNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKGPKAT 540
DB 481 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKGPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCPAEKGGKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDKETCPAEKGGKLVAAASQAALGL 585

RESULT 6
US-09-322-322-445
; Sequence 445, Application US/0932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Beltzer, James P.
; APPLICANT: Pottier, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/0932,322
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
```

```
; ORGANISM: Homo Sapiens
US-09-932-322-445

Query Match      100.0%; Score 3103; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60
DB 1 DAHSEVAHFRKDLGEENFKALVLIAPAOYLQOCPPFEDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFDGKLCITVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFDGKLCITVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNBEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHDNBEETFLKYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPILLESKHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKCCCKPILLESKHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360
DB 301 DLPSLAADPVESKDVCKNYAEAKOVFLGMFLYEVARRHPDYSVVLRLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVFDEPKPLVEEPQNLIKONCELFEQLGKGFQNALLVRYTKVPQVST 420
DB 361 CAADAPHECYAKVFDEPKPLVEEPQNLIKONCELFEQLGKGFQNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPKAEKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKGPKAT 540
DB 481 LVNRRPCPSALEVDETVVPKEFNAETFTFHADICTLSEKERQIKKQTALVELVKGPKAT 540
QY 541 KEQLKAVMDDFAAFVEKCKCKADDKETCPAEKGGKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAFVEKCKCKADDKETCPAEKGGKLVAAASQAALGL 585

RESULT 7
US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US2003019043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS42
; CURRENT APPLICATION NUMBER: US/09/832,501
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
```

FILING DATE: 19-DEC-1996

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 585 amino acids

TYPE: amino acid

STRANDEDNESS: <unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-984-010-26

Query Match

Best Local Similarity 100.0%; Score 3103; DB 10; Length 585;

Mismatches 0; Indels 0; Gaps 0;

Matches 585; Conservative 0;

QY

1

DAHKSVAHRFKDGLGEENFKALVLIAPAQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE

60

DB

1

DAHKSVAHRFKDGLGEENFKALVLIAPAQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE

60

QY

61

NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

DB

61

NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

QY

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

DB

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

QY

181

KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

DB

181

KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

QY

241

VHTECCGHDLLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA

300

DB

241

VHTECCGHDLLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA

300

QY

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

DB

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

QY

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST

420

DB

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST

420

QY

421

PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES

480

DB

421

PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES

480

QY

481

LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKERQIKKQTALVELVVKHPKAT

540

DB

481

LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKERQIKKQTALVELVVKHPKAT

540

QY

541

KEQLKAVNDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

DB

541

KEQLKAVNDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

RESULT 4

US-09-833-041-18

Sequence 18, Application US/09833041

Publication No. US20030125247A1

GENERAL INFORMATION:

APPLICANT: Haseltine, William A.

APPLICANT: Rosen, Craig A.

TITLE OF INVENTION: Albumin Fusion Proteins

FILE REFERENCE: PF545

CURRENT APPLICATION NUMBER: US/09/833,041

CURRENT FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: 60/229,358

PRIOR FILING DATE: 2000-04-12

PRIOR APPLICATION NUMBER: 60/256,931

PRIOR FILING DATE: 2000-12-21

QY

1

DAHKSVAHRFKDGLGEENFKALVLIAPAQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE

60

DB

1

DAHKSVAHRFKDGLGEENFKALVLIAPAQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE

60

QY

61

NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

DB

61

NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

QY

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

DB

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

QY

181

KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

DB

181

KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

QY

241

VHTECCGHDLLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA

300

DB

241

VHTECCGHDLLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA

300

QY

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

DB

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

QY

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST

420

DB

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST

420

QY

421

PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES

480

DB

421

PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES

480

QY

481

LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKERQIKKQTALVELVVKHPKAT

540

DB

481

LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKERQIKKQTALVELVVKHPKAT

540

QY

541

KEQLKAVNDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

DB

541

KEQLKAVNDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

RESULT 5

US-09-833-117-18

Sequence 18, Application US/09833117

Publication No. US20030171267A1

GENERAL INFORMATION:

APPLICANT: Rosen, Craig A.

APPLICANT: Sadeghi, Homa

APPLICANT: Prior, Christopher P.

APPLICANT: Turner, Andrew J.

TITLE OF INVENTION: Albumin Fusion Proteins

FILE REFERENCE: PF543

CURRENT APPLICATION NUMBER: US/09/833,117

CURRENT FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: 60/229,358

PRIOR FILING DATE: 2000-04-12

PRIOR APPLICATION NUMBER: 60/256,931

PRIOR FILING DATE: 2000-12-21

QY

1

DAHKSVAHRFKDGLGEENFKALVLIAPAQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE

60

DB

1

DAHKSVAHRFKDGLGEENFKALVLIAPAQYLOQCPEHDHVKLVNEVTEFAKTCVADESAAE

60

QY

61

NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

DB

61

NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV

120

QY

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

DB

121

DVMCTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP

180

QY

181

KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

DB

181

KLDELDEGKASSAKQRLKCSLOKFGERAFKAWAVARLSORPPKAEFAEVSCLVTDLT

240

QY

241

VHTECCGHDLLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA

300

DB

241

VHTECCGHDLLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA

300

QY

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

DB

301

DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTTLEKC

360

QY

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST

420

DB

361

CAAADPHECYAKVDFEFPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST

420

QY

421

PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES

480

DB

421

PTLVEVSRLNGKVGSKCKKHPKAEKMPCAEDYLSVNLQCLVLEHKTTPVSDRVTKCCTES

480

QY

481

LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKERQIKKQTALVELVVKHPKAT

540

DB

481

LVNRRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKERQIKKQTALVELVVKHPKAT

540

QY

541

KEQLKAVNDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

DB

541

KEQLKAVNDDPAAAFVEKCKCKADDKETCFABEGKKLVAAASQAALGL

585

```

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2

Query Match      100.0%; Score 3103; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. NO. 4.2e-269;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGSENFKALVLIAPAOVLOQCPEDHVKLVNEVTEFAKTCVADESSE 60
DB 1 DAHSEVAHRFKDGLGSENFKALVLIAPAOVLOQCPEDHVKLVNEVTEFAKTCVADESSE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
DB 121 DVNCTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP 180
QY 181 KLDELDEGRKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLTG 240
DB 181 KLDELDEGRKASSAKQRLKASLQKGERAFKAWAVARLSQRPFKAEFAEVSCLVTDLTG 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKHCIAEVNDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLKECCCKPILLESKHCIAEVNDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVVLRLRLAKTYETITLTK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVVLRLRLAKTYETITLTK 360
QY 361 CAADPHCEYAKVDFEPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
DB 361 CAADPHCEYAKVDFEPKPLVEEPQNLIKONCELFQOLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCABDYLVSVLNQLCVLHEKTPVSDRVTCKCTES 480
DB 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCABDYLVSVLNQLCVLHEKTPVSDRVTCKCTES 480
QY 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
DB 481 LVNRRPCFSALEVDVETVPKEFNAETFTFHADICTLSEKERQIKQTALVELVKHKPKAT 540
QY 541 KEQLKAWMDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAWMDPFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585

RESULT 3
US-09-984-010-26
; Sequence 26, Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 246.316 Seconds
(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18
Perfect score: 3103
Sequence: 1 DAKSEVAHRFDLGEENFK.....TCFAEGKKLVASQAALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PTC_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	9	US-09-929-552-2
2	3103	100.0	585	10	US-09-932-613-445
3	3103	100.0	585	10	US-09-984-010-26
4	3103	100.0	585	10	US-09-833-041-18
5	3103	100.0	585	10	US-09-833-117-18
6	3103	100.0	585	10	US-09-932-322-445
7	3103	100.0	585	10	US-09-832-501-18
8	3103	100.0	585	11	US-09-833-118-18
9	3103	100.0	585	11	US-09-833-245-18
10	3103	100.0	585	12	US-10-424-999-11
11	3103	100.0	585	12	US-10-425-000-31
12	3103	100.0	585	13	US-10-153-064-5
13	3103	100.0	585	14	US-10-153-604A-5
14	3103	100.0	585	14	US-10-319-263-1
15	3103	100.0	585	14	US-10-319-263-2

16	3103	100.0	585	14	US-10-414-469-1	Sequence 1, Appli
17	3103	100.0	585	14	US-10-414-469-2	Sequence 2, Appli
18	3103	100.0	585	14	US-10-413-831-1	Sequence 1, Appli
19	3103	100.0	585	14	US-10-413-831-2	Sequence 2, Appli
20	3103	100.0	585	15	US-10-413-832-1	Sequence 1, Appli
21	3103	100.0	585	15	US-10-413-832-2	Sequence 2, Appli
22	3103	100.0	585	15	US-10-414-386-1	Sequence 1, Appli
23	3103	100.0	585	15	US-10-414-386-2	Sequence 2, Appli
24	3103	100.0	585	15	US-10-233-675A-11	Sequence 11, Appl
25	3103	100.0	585	15	US-10-462-262-26	Sequence 26, Appl
26	3103	100.0	604	10	US-09-984-010-7	Sequence 7, Appli
27	3103	100.0	609	10	US-09-919-039-370	Sequence 370, App
28	3103	100.0	609	13	US-10-153-064-7	Sequence 7, Appli
29	3103	100.0	609	14	US-10-153-604A-7	Sequence 7, Appli
30	3103	100.0	610	9	US-09-984-186-2	Sequence 23, Appl
31	3103	100.0	610	14	US-10-237-667-2	Sequence 2, Appli
32	3103	100.0	610	14	US-10-237-708-2	Sequence 2, Appli
33	3103	100.0	610	14	US-10-237-866-2	Sequence 2, Appli
34	3103	100.0	610	14	US-10-237-871-2	Sequence 2, Appli
35	3103	100.0	610	14	US-10-237-871-2	Sequence 2, Appli
36	3103	100.0	610	14	US-10-237-624-2	Sequence 13, Appl
37	3103	100.0	616	12	US-10-433-108-13	Sequence 16, Appl
38	3103	100.0	624	12	US-10-433-108-16	Sequence 14, Appl
39	3103	100.0	631	12	US-10-433-108-14	Sequence 15, Appl
40	3103	100.0	640	12	US-10-433-108-15	Sequence 17, Appl
41	3103	100.0	640	12	US-10-433-108-17	Sequence 133, App
42	3103	100.0	651	13	US-10-153-064-133	Sequence 133, App
43	3103	100.0	651	14	US-10-153-604A-133	Sequence 132, App
44	3103	100.0	652	13	US-10-153-064-132	Sequence 132, App
45	3103	100.0	652	14	US-10-153-604A-132	Sequence 132, App

ALIGNMENTS

RESULT 1
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnschein, Carlos
; Scto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids

QY 421 PTLVEVRNLGKVGSKCCCHPEAKMPKPCABDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVRNLGKVGSKCCCHPEAKMPKPCABDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKEROIKKQTALVELVKKPKAT 540
 DB 481 LVNRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKEROIKKQTALVELVKKPKAT 540
 QY 541 KEQLKAVMDPFAAFVEKCKCKADDDKCTCFABEGKGLVAASQAALGL 585
 DB 541 KEQLKAVMDPFAAFVEKCKCKADDDKCTCFABEGKGLVAASQAALGL 585

RESULT 15
 AB-00986
 ID ABJ00986 standard; protein; 585 AA.
 AC ABJ00986;
 DT 05-SEP-2002 (first entry)
 XX B lymphocyte stimulator protein binding peptide related protein.
 DE B lymphocyte stimulator protein binding protein; BlyS; immune disease;
 KW allergy; proliferative disease; infectious disease; arteriosclerosis;
 KW inflammatory disorder; hypergammaglobulinaemia; blood clotting;
 KW ischaemia; graft-versus-host disease; neurodegenerative disease;
 KW immunosuppressive; nephrotropic; antithrombotic; antiarthritic;
 KW neuroprotective; cytostatic; immunostimulant; antitumour; anti-HIV;
 KW antidiabetic; antiallergic; thyromimetic; antinaemic; haemostatic;
 KW dermatological; antiinflammatory; cardiant; ophthalmological; uropathic;
 KW antidiabetic; antithyroid; antidepressant; hepatotropic.
 XX Homo sapiens.
 OS
 XX
 XX W0200216411-A2.
 XX
 XX 28-FEB-2002.
 XX
 XX 17-AUG-2001; 2001WO-US025850.
 XX
 XX 18-AUG-2000; 2000US-0226700P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Beltzer JP, Potter DM, Fleming TL, Rosen CA;
 XX
 XX WPI; 2002-499775/53.
 XX
 XX The treatment of various diseases e.g. rheumatoid arthritis, comprises
 PT administering B Lymphocyte stimulator binding polypeptide.
 PT
 XX Disclosure; Page 379-382; 387pp; English.

XX The present invention relates to the treatment, prevention or
 CC amelioration of a disease or disorder associated with: aberrant B
 CC lymphocyte stimulator (BlyS), BlyS receptor expression or activity; cells
 CC of haematopoietic origin; or proliferative disease; and reducing
 CC inhibiting or stimulating immunoglobulin production, B cell proliferation
 CC and graft rejection involving administration of BlyS binding polypeptide.
 CC The BlyS binding polypeptides are used in the treatment, prevention or
 CC amelioration of diseases such as immune system diseases, proliferative
 CC diseases, diseases of cells of hematopoietic origin, graft rejection,
 CC allergies, infectious diseases, arteriosclerosis, inflammatory disorders,
 CC hypergammaglobulinaemia, blood clotting disorders, ischaemia, and
 CC neurodegenerative diseases. The present sequence is a protein described
 CC in the invention
 XX
 XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 5; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKSEVAHRFKDLGSENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAB 60
 DB 1 DAHKSEVAHRFKDLGSENFKALVLIAPAOYLQCCPEDHVKLVNEVTEFAKTCVADESAB 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 120
 QY 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFVAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
 DB 121 DVNCTAFHDNEETFLKKYLYEIAARRHPYFVAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
 QY 181 KLDELDRDEGRKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
 DB 181 KLDELDRDEGRKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSKLVTDLTK 240
 QY 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADDDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKVYAEAKDVFLGMFLYEAARRHPDYSVVLRLAKTYETTELK 360
 DB 301 DLPSLAADFVESKDVCKVYAEAKDVFLGMFLYEAARRHPDYSVVLRLAKTYETTELK 360
 QY 361 CAAADPHECVAKVPDEFKPLVBEFQNLIKQNCLEFQOLGEYKFNALLVRYTKVQVST 420
 DB 361 CAAADPHECVAKVPDEFKPLVBEFQNLIKQNCLEFQOLGEYKFNALLVRYTKVQVST 420
 QY 421 PTLVEVRNLGKVGSKCCCHPEAKMPKPCABDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVRNLGKVGSKCCCHPEAKMPKPCABDYLSVNLQCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKEROIKKQTALVELVKKPKAT 540
 DB 481 LVNRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKEROIKKQTALVELVKKPKAT 540
 QY 541 KEQLKAVMDPFAAFVEKCKCKADDDKCTCFABEGKGLVAASQAALGL 585
 DB 541 KEQLKAVMDPFAAFVEKCKCKADDDKCTCFABEGKGLVAASQAALGL 585

Search completed: April 19, 2004, 11:51:14
 Job time : 348.086 secs

Matches	585;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	DAKSEVAHRFKDGBENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESAAE	60						
Db	1	DAKSEVAHRFKDGBENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESAAE	60						
QY	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV	120						
Db	61	NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV	120						
QY	121	DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP	180						
Db	121	DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP	180						
QY	181	KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLT	240						
Db	181	KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLT	240						
QY	241	VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA	300						
Db	241	VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA	300						
QY	301	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYVSVLLRLAKTYETTLK	360						
Db	301	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYVSVLLRLAKTYETTLK	360						
QY	361	CAAADPHECYAKVDFDEPKLVEBPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST	420						
Db	361	CAAADPHECYAKVDFDEPKLVEBPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST	420						
QY	421	PTLVEVSRNLGKVGSKCCHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCCCTES	480						
Db	421	PTLVEVSRNLGKVGSKCCHPKAPKMPCAEDYLSVVLNQLCVLHEKTPVSDRVTCCCTES	480						
QY	481	LVNRRFCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVKKHPRAT	540						
Db	481	LVNRRFCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQOTALVELVKKHPRAT	540						
QY	541	KQQLKAVMDDPAFVBEKCKADDKETCFABEGKGLVAASQAALGL	585						
Db	541	KQQLKAVMDDPAFVBEKCKADDKETCFABEGKGLVAASQAALGL	585						

RESULT 14

AAU75220
ID AAU75220 standard; protein; 585 AA.

XX AAU75220;

AC AAU75220;

DT 21-MAY-2002 (first entry)

XX Mature form of human serum albumin (HSA or HA).

DE Albumin fusion protein; therapeutic protein; immune disorder;

XX autoimmune disorder; blood-related disorder; hyperproliferative disorder;

KW renal disorder; cardiovascular disorder; respiratory disorder;

KW neurological disorder; endocrine disorder; reproductive system disorder;

KW gastrointestinal disorder; infectious disease; wound healing;

KW human serum albumin; HSA; HA.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Protein 1..585

FT /label= Mature_HSA

XX

PN WO200179271-A1.

XX

XX 25-OCT-2001.

XX

XX 12-APR-2001; 2001WO-US012009.

XX

XX 12-APR-2000; 2000US-0229358P.

PR

PR 25-APR-2000; 2000US-0199384P.
PR 21-DEC-2000; 2000US-0256931P.

PA (PRIN-) PRINCIPIA PHARM CORP.

PA (DELZ) DELTA BIOTECHNOLOGY LTD.

XX

PI Ballance DJ, Sleep D, Turner AJ, Sadeghi H, Prior CP;

XX WPI; 2002-179329/23.

DR N-PSDB; ABK13862.

XX New albumin fusion proteins with extended shelf life, useful for treating

PT leukemia, warts, hepatitis, multiple sclerosis and AIDS, comprises

PT therapeutic protein fused to albumin.

XX Claim 1; Fig 15; 338pp; English.

XX The present invention relates to albumin fusion proteins comprising a

CC therapeutic protein and albumin (e.g. human serum albumin (HSA or HA)).

CC The albumin fusion proteins are useful for treating, preventing, or

CC ameliorating various disorders. Such disorders include immune disorders,

CC autoimmune disorders, blood-related disorders, hyperproliferative

CC disorders, renal disorders, cardiovascular disorders, respiratory

CC disorders, neurological disorders, endocrine disorders, reproductive

CC system disorders, gastrointestinal disorders, infectious disease, and

CC wound healing. Therapeutic proteins can be stabilised to extend shelf

CC life and/or retain the protein's activity for extended periods of time in

CC solution, in vivo or in vitro by genetically or chemically fusing the

CC protein to albumin or its fragment or variant. In addition the use of

CC albumin fusion proteins reduces the need to formulate protein solutions

CC with large excesses of carrier proteins to prevent loss of therapeutic

CC protein due to factors such as binding to the container. The extension of

CC shelf life was tested by measuring biological activity (Nb2 cell

CC proliferation) of human albumin-human growth hormone (HA-hGH) fusion

CC protein remaining after incubation in cell culture media for up to 3

CC weeks. At week 3 there was still approximately 95% cell proliferation

CC compared to no activity of unfused hGH. The present sequence represents

CC the mature form of HSA which can be used to produce the albumin fusion

CC proteins of the invention

XX

SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 5; Length 585;

Best Local Similarity 100.0%; Pred. No. 1.5e-254;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFKOLGEENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60

Db 1 DAKSEVAHRFKOLGEENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120

Db 61 NCCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120

QY 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

Db 121 DVMCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLT 240

Db 181 KLDELDEGKASSAKQRLKASLOKFGERAFAKAWAVARLSQRPFAEPAEVSCLVTDLT 240

QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

Db 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYVSVLLRLAKTYETTLK 360

Db 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEAARRHPDYVSVLLRLAKTYETTLK 360

QY 361 CAAADPHECYAKVDFDEPKLVEBPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 420

Db 361 CAAADPHECYAKVDFDEPKLVEBPQNLIKQNCLEFQELGEYKFNQALLVRYTKVPQVST 420

FT Disulfide-bond 437.448
FT Disulfide-bond 461.477
FT Disulfide-bond 476.487
FT Region 492.511
FT /note="flexible inter-subdomain linker region"
FT 512.585
FT /label="subdomain"
FT Disulfide-bond 514.559
FT Disulfide-bond 558.567
XX WO200179442-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US011850.
XX
XX 12-APR-2000; 2000US-0229358P.
XX 25-APR-2000; 2000US-0199384P.
XX 21-DEC-2000; 2000US-0256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX Rosen CA, Haseltine WA;
XX WPI; 2001-611723/70.
XX N-PSDB; ABN87288.
XX
XX
XX New albumin fusion proteins, useful for treating diseases and disorders
XX such as cancer, comprise therapeutic protein fused to albumin.
XX
XX Claim 1; Fig 11; 413pp; English.
XX
XX The present invention describes an albumin fusion protein (I) comprising
XX a therapeutic protein: X and (a fragment or variant of) albumin
XX comprising a the fully defined sequence in AB879006 of 585 amino acids,
XX (where the fragment or variant has albumin or therapeutic protein: X
XX activity). (I) can have cytostatic, anorectic, immunosuppressive,
XX antidiabetic, antineumatic, antiarthritic and psoriatic activities.
XX Albumin fusion proteins are stabilised therapeutic proteins e.g.
XX antibodies to C5, C242 and CD80 useful for treating various diseases and
XX disorders such as non-Hodgkin's lymphoma, cancer, obesity, transplant
XX rejection, type I diabetes mellitus, rheumatoid arthritis and psoriasis.
XX Fusing albumin to therapeutic proteins stabilises the therapeutic
XX protein, extends the shelf life and retains the in vitro or in vivo
XX biological activity. It also reduces the need to formulate protein
XX solutions with large excesses of carrier proteins to prevent loss of
XX therapeutic proteins due to factors such as binding to the container. The
XX fusion proteins are easily dispensed with a simple formulation requiring
XX minimal post storage manipulation. The fusion of therapeutic proteins to
XX albumin confers stability in aqueous or other solution. The present
XX sequence represents the mature human albumin (HA) protein which is used
XX in the exemplification of the present invention
XX
XX Sequence 585 AA;
XX
XX Query Match 100.0%; Score 3103; DB 4; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-254;
XX Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSEVAHRFKDIGNENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHKSEVAHRFKDIGNENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNENLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNENLRLVRPEV 120
QY 121 DVMCTAFHDNEETLKKLYEIAARRHPFYFAPPELLFPFAKRYKAAPTECCQAADKAACLLP 180
DB 121 DVMCTAFHDNEETLKKLYEIAARRHPFYFAPPELLFPFAKRYKAAPTECCQAADKAACLLP 180
QY 181 KLDELREGKASSAKQRLKCSLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240

DB 181 KLDELREGKASSAKQRLKCSLOKFGERAFKAWAVARLSQRPKAEFAEVSKLVTDLTK 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCECKPPLLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCECKPPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLKCC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVVLRLAKTYETTLKCC 360
QY 361 CAAADPHCEYAKVFDEFKPLVEEPQNLIKONCELFQELGEYKFQNALLVRYTKKVPQVST 420
DB 361 CAAADPHCEYAKVFDEFKPLVEEPQNLIKONCELFQELGEYKFQNALLVRYTKKVPQVST 420
QY 421 PTLVEVSRNLKGVSKCKCKPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRNLKGVSKCKCKPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALVEDETVVPKEFNAETTFHADICTLSEKERQIKKOTALVELVXHKPKAT 540
DB 481 LVNRRPCFSALVEDETVVPKEFNAETTFHADICTLSEKERQIKKOTALVELVXHKPKAT 540
QY 541 KEQLKAVMDDPAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 541 KEQLKAVMDDPAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
XX
XX
XX RESULT 13
XX AAE08578 standard; protein; 585 AA.
XX
XX AAE08578;
XX
XX 19-NOV-2001 (first entry)
XX Human serum albumin (HSA).
XX
XX Human; albumin; cancer; cell proliferation; drug screening; biopsy.
XX Homo sapiens.
XX US6274305-B1.
XX
XX 14-AUG-2001.
XX
XX 19-DEC-1996; 96US-00769746.
XX
XX 19-DEC-1996; 96US-00769746.
XX (TUFT) UNIV TUFTS.
XX
XX Sonnenschein C, Soto AM;
XX
XX WPI; 2001-540371/60.
XX N-PSDB; AAD11488.
XX
XX Measuring human cell proliferation, useful in drug screening to determine
XX the potential for inhibiting cancer cell proliferation and for evaluating
XX biopsied tumors, comprises employing albumin-derived peptide.
XX
XX Claim 1; Fig 1; 20pp; English.
XX
XX The invention related to a method for testing cancer cells. The method is
XX useful for measuring human cancer cell proliferation, particularly for
XX determining the potential for inhibiting cancer cells proliferation using
XX albumin-derived peptides. The invention is also useful for drug screening
XX assays, as well as for evaluating biopsied tumors. The present sequence
XX is human serum albumin (HSA) related to the invention
XX
XX Sequence 585 AA;
XX
XX Query Match 100.0%; Score 3103; DB 4; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-254;

FT Domain /label= Loop_IX
FT 461..475
FT /label= Loop_X
FT 478..486
FT /label= Loop_XI
FT 560..566
FT /label= Loop_XII
PN WO200179258-A1.
XX
XX
PD 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US012008.
XX
XX 12-APR-2000; 2000US-0229359P.
PR 25-APR-2000; 2000US-0193384P.
PR 21-DEC-2000; 2000US-0256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (PRIN-) PRINCIPIA PHARM CORP.
XX
XX Rosen CA, Sadeghi H, Prior CP, Turner AJ;
XX
XX WPI; 2001-602931/68.
DR N-PSDB; AAD22287.
XX
XX Albumin fusion proteins comprising a therapeutic protein and albumin,
PT useful in the treating metastatic renal cell carcinoma, metastatic
PT melanoma, malignant melanoma, renal cell carcinoma, HIV (human
PT immunodeficiency virus) or infection.
XX
XX Claim 1; Fig 9; 325pp; English.
XX
XX The invention relates to albumin fusion proteins comprising therapeutic
CC protein and human albumin (HA). The albumin fusion proteins are useful in
CC the treatment, prevention, diagnosis, and/or detection of diseases,
CC disorders such as immune system disorders (transplant rejection); blood
CC related disorders (myocardial infarction); hyperproliferative disorders
CC (childhood acute myeloid leukaemia); renal disorder (glomerulonephritis);
CC cardiovascular disorders (arrhythmias); respiratory disorders (non-
CC allergic rhinitis); neurological diseases (Alzheimer's disease);
CC endocrine disorders (pheochromocytoma); reproductive system disorders
CC (syphilis); infectious diseases (measles); gastrointestinal disorders
CC (irritable bowel syndrome) and wound healing. The albumin fusion proteins
CC are also used in the treatment of metastatic renal cell carcinoma,
CC metastatic melanoma, malignant melanoma and HIV (human immunodeficiency
CC virus) infection. Nucleic acid encoding albumin fusion protein is useful
CC in gene therapy. The present sequence is human albumin (HA) protein
XX
XX Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHKSVAHRFKDLGEENFKALVLIAPQVLOCCPEDHVKLVNEVTEPAKTCVADESAG 60
DB 1 DAHKSVAHRFKDLGEENFKALVLIAPQVLOCCPEDHVKLVNEVTEPAKTCVADESAG 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPLRVPRPV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDPNLPLRVPRPV 120
QY 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFPKRYKAAFTCCQAAKACLLP 180
DB 121 DVNCTAFHNEETFLKKLYEIAARRHPYFYAPPELLFPKRYKAAFTCCQAAKACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPKAEFAVSKLVTDLTK 240
QY 241 VHTCCGHDLLFCADDEADLAKYICENQDSISSKLKCECKPFLLEKSHCIAEVENDEMPA 300

Db 241 VHTCCGHDLLFCADDEADLAKYICENQDSISSKLKCECKPFLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAEARRHPDYVLLLRLLAKTYETLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYAEARRHPDYVLLLRLLAKTYETLEKC 360
QY 361 CAADDPHECVAKVDFDEKPLVEEPQNLIKQNCLEPEQLGEYKFNQALLVRYTKVPQVST 420
Db 361 CAADDPHECVAKVDFDEKPLVEEPQNLIKQNCLEPEQLGEYKFNQALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLKGVSKCKCKHPEAKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLKGVSKCKCKHPEAKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPKAT 540
Db 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKEROIKKOTALVELVKKHPKAT 540
QY 541 KEOLKAVMDDDFAAFVEKCKKADDKETCFPAEEGKKLVAAASQAALGL 585
Db 541 KEOLKAVMDDDFAAFVEKCKKADDKETCFPAEEGKKLVAAASQAALGL 585
RESULT 12
ABB79006
ID ABB79006 standard; protein; 585 AA.
XX
XX ABB79006;
XX
XX 01-AUG-2002 (first entry)
XX
XX Human mature albumin protein SEQ ID NO:18.
XX
XX Human; growth hormone; hGH; albumin; human serum albumin; HSA;
KW albumin fusion protein; cytostatic; anorectic; immunosuppressive;
KW antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;
KW non-Hodgkin's lymphoma; obesity; transplant rejection; psoriasis;
KW type I diabetes mellitus; rheumatoid arthritis.
XX
XX Homo sapiens.
FH Key Location/Qualifiers
FT Domain 1..194
FT /label= 1
FT Domain 1..105
FT /label= subdomain
FT Disulfide-bond 53..62
FT Disulfide-bond 75..91
FT Disulfide-bond 90..101
FT Region 106..119
FT /note= "flexible inter-subdomain linker region"
FT Domain 120..194
FT /label= subdomain
FT Disulfide-bond 124..169
FT Disulfide-bond 168..177
FT Domain 195..387
FT /label= 2
FT Domain 195..291
FT /label= subdomain
FT Disulfide-bond 245..253
FT Disulfide-bond 265..279
FT Disulfide-bond 278..289
FT Region 292..315
FT /note= "flexible inter-subdomain linker region"
FT Domain 316..387
FT /label= subdomain
FT Disulfide-bond 316..361
FT Disulfide-bond 360..369
FT Domain 388..585
FT /label= 3
FT Domain 388..491
FT /label= subdomain
FT Disulfide-bond 392..438

FT	Domain	/label= Loop_III	
FT		170..176	
FT		/label= Loop_IV	
FT		247..252	
FT	Domain	/label= Loop_V	
FT		266..277	
FT		/label= Loop_VI	
FT		280..288	
FT	Domain	/label= Loop_VII	
FT		362..368	
FT		/label= Loop_VIII	
FT		439..447	
FT	Domain	/label= Loop_IX	
FT		461..475	
FT		/label= Loop_X	
FT		478..485	
FT	Domain	/label= Loop_XI	
FT		560..566	
FT		/label= Loop_XII	
XX	WC200179443-A2.		
XX	25-OCT-2001.		
XX	12-APR-2001; 2001WO-US011924.		
XX	12-APR-2000; 2000US-0229358P.		
XX	25-APR-2000; 2000US-0199384P.		
XX	21-DEC-2000; 2000US-0256931P.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
XX	Rosen CA, Haseltine WA;		
XX	WPI; 2001-616754/71.		
XX	N-PSDB; AAD21639.		
XX	Albumin fusion proteins comprising a therapeutic protein and albumin,		
XX	useful in the treating immune system disorders (e.g. transplant		
XX	rejection), blood related disorders (e.g. myocardial infarction) and		
XX	hyperproliferative disorders.		
XX	Claim 1; Fig 9; 380pp; English.		
XX	The invention relates to albumin fusion proteins comprising therapeutic		
XX	protein and human albumin (HA). Therapeutic protein fused to albumin have		
XX	an extended shelf-life. The albumin fusion proteins are useful in the		
XX	treatment, prevention, diagnosis and/or detection of diseases, disorders		
XX	such as immune system disorders (e.g. transplant rejection), blood		
XX	related disorders (e.g. myocardial infarction), hyperproliferative		
XX	disorders (e.g. childhood acute myeloid leukaemia), renal disorders (e.g.		
XX	glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),		
XX	respiratory disorders (e.g. non-allergic rhinitis), neurological diseases		
XX	(e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma),		
XX	reproductive system disorders (e.g. syphilis), infectious diseases (e.g.		
XX	measles), gastrointestinal disorders (e.g. irritable bowel syndrome) and		
XX	wound healing. Nucleic acids encoding albumin fusion protein is used in		
XX	gene therapy. The present sequence is human albumin (HA) protein		
XX	Sequence 585 AA;		
SQ	Query Match	100.0%; Score 3103; DB 4; Length 585;	
	Best Local Similarity	100.0%; Pred. No. 1.5e-254;	
	Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	DAHSEVAHRFKDGEENFKALVIAFAQYLCQCFEDHVKLVNEVTEFAKTCVADESA 60	
Db	1	DAHSEVAHRFKDGEENFKALVIAFAQYLCQCFEDHVKLVNEVTEFAKTCVADESA 60	
Qy	61	NCDSKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120	
Db	61	NCDSKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120	

Qy	121	DVMCTAFHDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP	180
Db	121	DVMCTAFHDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAACLLP	180
Qy	181	KLDELDEGKASSAKORLKCASLQKFGERAFAWAVARLSORPPKAEFAEVSCLVTDLT	240
Db	181	KLDELDEGKASSAKORLKCASLQKFGERAFAWAVARLSORPPKAEFAEVSCLVTDLT	240
Qy	241	VHTECCHGDLLECCADDRADLAKYICENQDSISSKLECCCEKPLEKSHCIAEVENDEMPA	300
Db	241	VHTECCHGDLLECCADDRADLAKYICENQDSISSKLECCCEKPLEKSHCIAEVENDEMPA	300
Qy	301	DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYIYARRHPDYSVLLLRLLAKTYETTLEK	360
Db	301	DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYIYARRHPDYSVLLLRLLAKTYETTLEK	360
Qy	361	CAAADPHECYAKVDFEKFPLVEBPQNLKONCELPQOLGEYKFNQALLVRYTKKVPQVST	420
Db	361	CAAADPHECYAKVDFEKFPLVEBPQNLKONCELPQOLGEYKFNQALLVRYTKKVPQVST	420
Qy	421	PTLVEVSRNLGKVGSKCKHPEAKRMPACADYLSVVLNQLCVLHEKTPVSDRYTKCCTES	480
Db	421	PTLVEVSRNLGKVGSKCKHPEAKRMPACADYLSVVLNQLCVLHEKTPVSDRYTKCCTES	480
Qy	481	LVNERPCFSALVEYDVTVPKEFNAETFTHADICTLSEKERQIKKOTALVELVKHKPKAT	540
Db	481	LVNERPCFSALVEYDVTVPKEFNAETFTHADICTLSEKERQIKKOTALVELVKHKPKAT	540
Qy	541	KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL	585
Db	541	KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKKLVAAASQAALGL	585
RESULT 11			
AAE13399			
ID	AAE13399	standard; protein; 585 AA.	
XX	AC	AAE13399;	
XX	DT	12-FEB-2002 (first entry)	
XX	DE	Human albumin (HA) protein.	
XX	KW	Human; albumin; HA; fusion protein; immune system disorder; syphilis;	
XX	KW	transplant rejection; blood related disorder; myocardial infarction;	
XX	KW	hyperproliferative disorder; acute myeloid leukaemia; renal disorder;	
XX	KW	glomerulonephritis; cardiovascular disorder; arrhythmia; rhinitis;	
XX	KW	respiratory disorder; neurological disease; Alzheimer's disease;	
XX	KW	endocrine disorder; pheochromocytoma; reproductive system disorder;	
XX	KW	measles; gastrointestinal disorder; irritable bowel syndrome; HIV;	
XX	KW	human immunodeficiency virus; wound healing; renal cell carcinoma;	
XX	OS	melanoma; gene therapy.	
XX	OS	Homo sapiens.	
XX	Key	Location/Qualifiers	
FT	Domain	/label= Loop_I	54..61
FT	Domain	/label= Loop_II	76..89
FT	Domain	/label= Loop_III	92..100
FT	Domain	/label= Loop_IV	170..176
FT	Domain	/label= Loop_V	247..252
FT	Domain	/label= Loop_VI	266..277
FT	Domain	/label= Loop_VII	280..288
FT	Domain	/label= Loop_VIII	362..368
FT	Domain	/label= Loop_VIII	439..447

XX Key Location/Qualifiers
FH Domain 54..61
FT /label= Loop_I
FT Domain 76..89
FT /label= Loop_II
FT Domain 92..100
FT /label= Loop_III
FT Domain 170..176
FT /label= Loop_IV
FT Domain 247..252
FT /label= Loop_V
FT Domain 266..277
FT /label= Loop_VI
FT Domain 280..288
FT /label= Loop_VII
FT Domain 362..368
FT /label= Loop_VIII
FT Domain 439..447
FT /label= Loop_IX
FT Domain 461..475
FT /label= Loop_X
FT Domain 478..486
FT /label= Loop_XI
FT Domain 560..566
FT /label= Loop_XII
XX WO200179480-A1.
XX 25-OCT-2001.
XX 12-APR-2001; 2001WO-US011991.
XX 12-APR-2000; 2000US-0229358P.
XX 25-APR-2000; 2000US-0199384P.
XX 21-DEC-2000; 2000US-0256931P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Haseltine WA;
XX WPI; 2001-616756/71.
XX N-PSDB; AAD20005.
XX Albumin fusion proteins comprising a therapeutic protein and albumin,
XX useful in the treating metastatic renal cell carcinoma, metastatic
XX melanoma, malignant melanoma, renal cell carcinoma, HIV (human
XX immunodeficiency virus) or infection.
XX Claim 1; Fig 9; 394pp; English.
XX The invention relates to human albumin (HA) fusion proteins and their
XX corresponding nucleic acid sequences. Therapeutic proteins fused to
XX albumin or its fragments have an extended shelf-life. The albumin fusion
XX proteins are useful in the treatment, prevention, diagnosis, and/or
XX detection of diseases, disorders such as immune system disorders (e.g.
XX transplant rejection), blood related disorders (e.g. myocardial
XX infarction), hyperproliferative disorders (e.g. childhood acute myeloid
XX leukaemia, metastatic renal cell carcinoma, metastatic melanoma,
XX malignant melanoma, renal cell carcinoma), renal disorders (e.g.
XX glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),
XX respiratory disorders (e.g. non-allergic rhinitis), neurological diseases
XX (e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma),
XX reproductive system disorders (e.g. syphilis), infectious diseases (e.g.
XX measles), gastrointestinal disorders (e.g. irritable bowel syndrome), HIV
XX (human immunodeficiency virus) infection and wound healing. Nucleic acids
XX encoding albumin fusion protein is used in gene therapy. The present
XX sequence is human albumin
XX Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5e-254;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DAHSEVAHPRKDLGEENFKALVLIAPAYLOQCPEPDHVKLVNEVTEFAKTCVADESAE 60
Db 1 DAHSEVAHPRKDLGEENFKALVLIAPAYLOQCPEPDHVKLVNEVTEFAKTCVADESAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQSPERNECFLOHKDDNPRLVRPEV 120
Qy 121 DVNCTAFHNDNEETFLKKYLYETARRHPYFAPPELLFFAKRYKAAFTCCQAADKACALLP 180
Db 121 DVNCTAFHNDNEETFLKKYLYETARRHPYFAPPELLFFAKRYKAAFTCCQAADKACALLP 180
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLKVTDLT 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFPKAEFAEVSCLKVTDLT 240
Qy 241 VHTCCGDLLECCADDRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVNDMPA 300
Db 241 VHTCCGDLLECCADDRADLAKYICENQDSISSKLECCCKPLLEKSHCHIAEVNDMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVYLLRLAKTYETTLKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLVEYARRHPDYSVYLLRLAKTYETTLKC 360
Qy 361 CAAADPHECYAKVDFEFLVPEPQNLIKQNCLEFQOLGEYKFNALLVRYTKVQVST 420
Db 361 CAAADPHECYAKVDFEFLVPEPQNLIKQNCLEFQOLGEYKFNALLVRYTKVQVST 420
Qy 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHETKTPVSDRVTKCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHETKTPVSDRVTKCTES 480
Qy 481 LVNRRPCFSALEVDYTVVPKFNATFTTHADICTLSEKERQIKKOTALVELVKKHPKAT 540
Db 481 LVNRRPCFSALEVDYTVVPKFNATFTTHADICTLSEKERQIKKOTALVELVKKHPKAT 540
Qy 541 KEQLKAVMDDFAAVFEKCKCKADDKETCFABEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAVFEKCKCKADDKETCFABEGKKLVAASQAALGL 585
RESULT 10
AAE13129
ID AAE13129 standard; protein; 585 AA.
XX AAE13129;
AC AAE13129;
XX
DT 28-JAN-2002 (first entry)
XX Human albumin (HA).
XX Human; albumin; HA; fusion protein; therapeutic protein; vulneryary;
XX immune system disorder; transplant rejection; blood related disorder;
XX myocardial infarction; hyperproliferative disorder; glomerulonephritis;
XX childhood acute myeloid leukemia; cardiovascular disorder; arrhythmia;
XX respiratory disorder; gene therapy; non-allergic rhinitis; neutropenic;
XX neurological disease; Alzheimer's disease; reproductive system disorder;
XX endocrine disorder; pheochromocytoma; infectious disease; antiarthritis;
XX measles; gastrointestinal disorder; irritable bowel syndrome; syphilis;
XX wound healing; antineoplastic; immunosuppressive; neuroprotective;
XX cardiant; cytostatic; antileukemic; antirheumatic; antimicrobial;
XX renal disorder.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Domain 54..61
FT /label= Loop_I
FT Domain 76..89
FT /label= Loop_II
FT Domain 92..100

QY 361 CAADPHCEYAKVDFEKPILVEBPOMLIKONCELPOLGEYKFNALLVRYTKVPQVST 420
 DB 361 CAADPHCEYAKVDFEKPILVEBPOMLIKONCELPOLGEYKFNALLVRYTKVPQVST 420
 QY 421 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRPCFSALEVDVETVPKFNAEFTPHADICTLSEKEROIKQTALVELVHKPKAT 540
 DB 481 LVNRPCFSALEVDVETVPKFNAEFTPHADICTLSEKEROIKQTALVELVHKPKAT 540
 QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585
 DB 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585

RESULT 8

AAW52567
 ID AAW52567 standard; protein; 585 AA.

AC AAW52567;

DT 05-FEB-2002 (first entry)

XX Mature human serum albumin.

XX Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiant;
 KW nontropic; neuroprotective; gene therapy; immune disorder; wound healing;
 KW hyperproliferative disorder; renal disorder; cardiovascular disorder;
 KW respiratory disorder; neurological disease; endocrine disorder;
 KW reproductive system disorder; infectious disease;
 KW gastrointestinal disorder.

XX Homo sapiens.

XX WO200179444-A2.

XX 25-OCT-2001.

XX 12-APR-2001; 2001WO-US012013.

XX 12-APR-2000; 2000US-0229358P.

XX 21-DEC-2000; 2000US-0199384P.

XX 21-DEC-2000; 2000US-0256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Haseltine WA;

XX WPI, 2001-616755/71.

XX N-PSDB; ABA03057.

XX Albumin fusion proteins comprising a therapeutic protein and albumin,
 PT useful in the treating immune system disorders (e.g. transplant
 PT rejection), blood related disorders (e.g. myocardial infarction) and
 PT hyperproliferative disorders.

XX Claim 1; Fig 15; 606pp; English.

XX The present invention relates to albumin fusion proteins, which comprise
 CC a therapeutic protein and albumin. The present sequence is the protein
 CC sequence for mature human serum albumin (HA), which was used to generate
 CC the fusion proteins of the present invention. The albumin fusion proteins
 CC are useful in the treatment, prevention, diagnosis, and/or detection of
 CC diseases/disorders such as immune system disorders (e.g. transplant
 CC rejection), blood related disorders (e.g. myocardial infarction),
 CC hyperproliferative disorders (e.g. childhood acute myeloid leukemia),
 CC renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
 CC arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),
 CC neurological diseases (e.g. Alzheimer's disease), endocrine disorders
 CC (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),

CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
 CC irritable bowel syndrome) and wound healing
 XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 4; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVARRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEPAKTCVADSSAE 60
 DB 1 DAHSEVARRFKDLGEENFKALVLIAPAYLQOCFFEDHVKLVNEVTEPAKTCVADSSAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGPERNECFLOHKDDNPMLPLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKGPERNECFLOHKDDNPMLPLVRPEV 120
 QY 121 DVNCTAFHDNEETFLKKYLYEYARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
 DB 121 DVNCTAFHDNEETFLKKYLYEYARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKMAVARLSORFPKAEPAFVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKMAVARLSORFPKAEPAFVSKLVTDLTK 240
 QY 241 VHTCCCHGDLLECCADDRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCCHGDLLECCADDRADLAKYICENODSISSKLKECCCKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKVCNKYAEAKDVFLGMFLYEVARRHPDYVSVLLLLAKTYETTLBKC 360
 DB 301 DLPSLAADFVESKVCNKYAEAKDVFLGMFLYEVARRHPDYVSVLLLLAKTYETTLBKC 360
 QY 361 CAADPHCEYAKVDFEKPILVEBPOMLIKONCELPOLGEYKFNALLVRYTKVPQVST 420
 DB 361 CAADPHCEYAKVDFEKPILVEBPOMLIKONCELPOLGEYKFNALLVRYTKVPQVST 420
 QY 421 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRPCFSALEVDVETVPKFNAEFTPHADICTLSEKEROIKQTALVELVHKPKAT 540
 DB 481 LVNRPCFSALEVDVETVPKFNAEFTPHADICTLSEKEROIKQTALVELVHKPKAT 540
 QY 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585
 DB 541 KEQLKAVMDDFAAFVEKCKKADDKETCFABEGKGLVAASQAALGL 585

RESULT 9

AAE12403
 ID AAE12403 standard; protein; 585 AA.

XX AAE12403;

XX 18-DEC-2001 (first entry)

XX Human albumin (HA).

XX Human; albumin; HA; immune system disorder; transplant rejection;
 KW blood related disorder; myocardial infarction; glomerulonephritis;
 KW hyperproliferative disorder; childhood acute myeloid leukaemia;
 KW renal cell carcinoma; cardiovascular disorder; vulnery; melanoma;
 KW arrhythmia; respiratory disorder; non-allergic rhinitis; antileukemic;
 KW neurological disease; Alzheimer's disease; endocrine disorder; syphilis;
 KW pheochromocytoma; reproductive system disorder; neuroprotective; syphilis;
 KW infectious disease; gastrointestinal disorder; wound healing; nontropic;
 KW irritable bowel syndrome; HIV; human immunodeficiency virus infection;
 KW cytotoxic; antiinflammatory; gene therapy; immunosuppressive; cardiant;
 KW antiarthritic; antirheumatic; renal disorder; antimicrobial.

XX Homo sapiens.

XX New method for the continuous detection of ischemic states comprises
 PT detecting and quantifying the existence of an alteration of the serum
 PT protein albumin.
 XX
 XX Disclosure; Page 97-100; 105pp; English.
 XX
 CC The present sequence represents human albumin protein. The specification
 CC describes a method for the continuous detection of ischemic states. The
 CC method comprises detecting and quantifying the existence of an alteration
 CC of the serum protein albumin. The method comprises contacting a
 CC biological sample containing albumin from the patient with an excess
 CC quantity of a metal ion salt, where the metal ion binds to the N-terminus
 CC of naturally occurring human albumin, to form a mixture containing bound
 CC metal ions and unbound metal ions, and then determining the amount of
 CC metal ions bound to the albumin N-terminus. The amount of bound metal
 CC ions is correlated to a known value to determine the occurrence or non-
 CC occurrence of an ischemic event. The methods are useful for detection of
 CC ischemic states. The methods are also useful for distinguishing
 CC peroperative ischemia from ischemia caused by, amongst other things,
 CC myocardial infarctions and progressive coronary artery disease
 XX
 SQ Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 3; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGSENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHKSEVAHRFKDLGSENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
 DB 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180

QY 181 KLDELDRDEGKASSAKORLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELDRDEGKASSAKORLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

QY 241 VHTTECHGDLLECADDRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300
 DB 241 VHTTECHGDLLECADDRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVVLLLRKAKTYETTLTK 360
 DB 301 DLPSSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVVLLLRKAKTYETTLTK 360

QY 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFEQLGEYKFNQALLVRYTKKVPQVST 420
 DB 361 CAAADPHECYAKVDFEFKPLVEEPQNLIKQNCLEFEQLGEYKFNQALLVRYTKKVPQVST 420

QY 421 PTLVEVSRNLGVGSKCCGHPKAMPKPCADYLSVLNQLCVLHKPTVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGVGSKCCGHPKAMPKPCADYLSVLNQLCVLHKPTVSDRVTKCCTES 480

QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVXHKPKAT 540
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKOTALVELVXHKPKAT 540

QY 541 KQGLKAVMDPFAAFVEKCKADDKETCFABEGKGLVAASQAALGL 585
 DB 541 KQGLKAVMDPFAAFVEKCKADDKETCFABEGKGLVAASQAALGL 585

RESULT 7

 AAY83946
 ID AAY83946 standard; protein; 585 AA.

XX

AAAY83946;
 28-JUL-2000 (first entry)
 Yeast codon-biased recombinant human serum albumin protein.
 Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
 overlapping oligonucleotide; expression vector.
 Homo sapiens.
 Synthetic.
 CNL239103-A.
 22-DEC-1999.
 17-JUN-1998; 98CN-00102506.
 17-JUN-1998; 98CN-00102506.
 (HAIJ-) HAIJI BIOENGINEERING CO LTD.
 Li S, Lu D;
 WPI; 2000-351198/31.
 N-PSDB; AAA10091.
 Process for preparing recombinant human serum albumin comprising yeast
 biased sex codons - uses a recombinant DNA technique.
 Disclosure; Fig 1; 44pp; Chinese.

The method relates to a method of recombinantly producing human serum
 albumin (HSA) in yeast by altering the coding sequence of HSA to comprise
 a yeast codon bias. The complete HSA gene (AAA10091) was generated as
 three synthetic fragments (AAA10092-AAA10094) joined by recombinant DNA
 technology. Each HSA fragment was synthesized from overlapping
 oligonucleotide fragments that were extended. This sequence represents
 the complete sequence of the HSA encoded by the human gene with a yeast
 codon bias. The invention also covers a recombinant expression vector,
 yeast host cells carrying the recombinant expression vector and the
 process for producing human serum albumin in the yeast host cell,
 especially in secretory mode

Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 3; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGSENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHKSEVAHRFKDLGSENFKALVLIAPAYQLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLTCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
 DB 121 DVNCTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180

QY 181 KLDELDRDEGKASSAKORLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240
 DB 181 KLDELDRDEGKASSAKORLKCSLQKFGERAFKAWAVARLSORFPKAEFAEVSKLVTDLTK 240

QY 241 VHTTECHGDLLECADDRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300
 DB 241 VHTTECHGDLLECADDRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300

QY 301 DLPSSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVVLLLRKAKTYETTLTK 360
 DB 301 DLPSSLAADFVESKDVCKNVAEAKDVFGLMFLYIARRHPDYSVVLLLRKAKTYETTLTK 360

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYETVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDYETVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540
QY 541 KEQLKAVMDDFAAFEVKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFEVKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 5
AAO20111
ID AAO20111 standard; protein; 585 AA.
AC AAO20111;
XX
XX
DT 06-AUG-2002 (first entry)
DE HSA protein sequence related to the growth hormone protein.
XX
XX
KW Serum albumin-growth hormone fusion protein; growth hormone;
KW Down's syndrome.
XX
XX
OS Unidentified.
XX
XX
PN KR99076789-A.
XX
XX
PD 15-OCT-1999.
XX
XX
PF 25-JUN-1998; 98KR-00704914.
XX
XX
PR 30-DEC-1995; 95GB-00026733.
PR 19-DEC-1996; 96WO-GB003164.
XX
XX
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
XX
PI Ballance DJ;
DR WPI; 1997-363680/33.
DR N-PSDB; AAK99568.
XX
XX
PT Serum albumin-growth hormone fusion protein - useful to treat growth
PT hormone related diseases, e.g. Down's syndrome.
XX
XX
PS Disclosure; Fig 6; 2lpp; Korean.
XX
XX
CC The invention relates to a serum albumin-growth hormone fusion protein -
CC useful to treat growth hormone related diseases such as Down's syndrome.
CC This sequence represents a HSA protein related to the serum albumin-
CC growth hormone protein of the invention
XX
SQ Sequence 585 AA;

Query Match
Best Local Similarity 100.0%; Score 3103; DB 2; Length 585;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDIDGENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAHFKDIDGENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVNCTAFHNEETFLKXLYIETARRHPYFVAPPELLFAFKRYKAAFTCCQOAKAACLLP 180
DB 121 DVNCTAFHNEETFLKXLYIETARRHPYFVAPPELLFAFKRYKAAFTCCQOAKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFPKAEFAEVSKLVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKASLQKFGERAFAKAWAVARLSQRPFPKAEFAEVSKLVTDLTK 240
QY 241 VHTCECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPELLEKSHCIAEVENDEMPA 300
DB 241 VHTCECHGDLLECCADRADLAKYICENQDSISSKLECCCKPPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNALLVRYTKVPQVST 420
DB 361 CAAADPHECYAKVDFEFKPLVEBPQNLIKONCELFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
DB 421 PTLVEVSRLGKVGSKCKKHPEAKMPCAEADYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYETVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540
DB 481 LVNRRPCFSALEVDYETVPKEFNAETFTPHADICTLSEKERQIKKQATLVELVGHKPKAT 540
QY 541 KEQLKAVMDDFAAFEVKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAFEVKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

RESULT 6
AA84873
ID AA84873 standard; protein; 585 AA.
AC AA84873;
XX
XX
DT 08-AUG-2000 (first entry)
DE Amino acid sequence of a human albumin protein.
XX
XX
KW Human; albumin; ischemic state; serum protein; metal ion salt;
KW perioperative ischemia; ischemia; myocardial infarction;
KW progressive coronary artery disease.
XX
XX
OS Homo sapiens.
XX
XX
PH Key Location/Qualifiers
FT Modified-site 1
FT /note= "optionally acetylated, and claimed under claim
FT 56"
XX
XX
PN WO200020840-A1.
XX
XX
PD 13-APR-2000.
XX
XX
PF 01-OCT-1999; 95WO-US022905.
XX
XX
PR 02-OCT-1998; 98US-00165581.
PR 02-OCT-1998; 98US-001655926.
PR 02-OCT-1998; 98US-0102738P.
PR 11-JAN-1999; 99US-0115392P.
XX
XX
PA (ISCH-) ISCHEMIA TECHNOLOGIES INC.
XX
XX
PI Bar-Or D, Lau E, Winkler JV;
XX WPI; 2000-303843/26.
DR

XX JP02227079-A.
 XX
 XX
 PD 10-SEP-1990.
 XX
 PF 25-AUG-1989; 89JP-00217540.
 XX
 XX 06-OCT-1988; 88JP-00250926.
 PR
 XX (TOFU) TONEN CORP.
 PA
 XX WPI; 1990-317325/42.
 DR N-PSDB; AAQ06059.
 XX
 XX New human serum albumin fragments - used to bond to medicines and for
 PT stable folding of protein(s).
 XX
 XX Claim 1; Fig 8; 24pp; Japanese.
 XX
 XX Fragments A-C of HSA are expressed as fusion proteins with the signal
 CC peptide of E. coli alkaline phosphatase. The fragments are selected for
 CC their specific properties. The C-terminal truncated fragment, B, does not
 CC bind long-chain fatty acids but does bind to various medicines at the
 CC central region. The N-terminal truncated fragment, C, has good stability
 CC in protein folding. The central segment, A, has characteristics of both B
 CC and C. See also AAQ06096-Q06098. (Updated on 25-MAR-2003 to correct PD
 CC field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-
 CC 2003 to correct PR field.)
 XX
 XX Sequence 585 AA;
 SQ
 Query Match 100.0%; Score 3103; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVMCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 121 DVMCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 QY 181 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVDTLTK 240
 DB 181 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVDTLTK 240
 QY 241 VHTECCHGDLLECADDDRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300
 DB 241 VHTECCHGDLLECADDDRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADPFVSKDVCKNVAEAKOVFLGMFLVEYARRHPDYSVLLLLAKTYETTLK 360
 DB 301 DLPSLAADPFVSKDVCKNVAEAKOVFLGMFLVEYARRHPDYSVLLLLAKTYETTLK 360
 QY 361 CAAADPHECVAKVDFEFKPLVEEPONLIKONCELFEOQLGEYKFONALLVRYTKVPOYST 420
 DB 361 CAAADPHECVAKVDFEFKPLVEEPONLIKONCELFEOQLGEYKFONALLVRYTKVPOYST 420
 QY 421 PTLVEVSRNLGVGSKCKKHPAKMPCAEADYLSVVLNQLCVLHEKTPVSRVTCKCTES 480
 DB 421 PTLVEVSRNLGVGSKCKKHPAKMPCAEADYLSVVLNQLCVLHEKTPVSRVTCKCTES 480
 QY 481 LVNRRPCFSALEVDETVYVPEFNAETFTPHADICTLSEKEROIKKOTALVELVHKPKAT 540
 DB 481 LVNRRPCFSALEVDETVYVPEFNAETFTPHADICTLSEKEROIKKOTALVELVHKPKAT 540
 QY 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585

DB 541 KEQLKAVMDPFAAFVEKCKCKADDKETCFABEGKCLVAASQAALGL 585
 RESULT 4
 AAR80301
 ID AAR80301 standard; protein; 585 AA.
 XX
 AC AAR80301;
 XX
 DT 25-MAR-2003 (revised)
 DT 17-JAN-1996 (first entry)
 XX
 DE Human serum albumin.
 XX
 KW Serum albumin; HSA; aspartyl protease-3; Yap3p; Saccharomyces cerevisiae.
 XX
 OS Homo sapiens.
 XX
 PN WO9523857-A1.
 XX
 PD 08-SEP-1995.
 XX
 PF 01-MAR-1995; 95WO-GB0000434.
 XX
 PR 05-MAR-1994; 94GB-00004270.
 XX
 PA (DELZ) DELTA BIOTECHNOLOGY LTD.
 XX
 PI KerryWilliams SM, Gilbert SC;
 XX
 DR WPI; 1995-320572/41.
 DR N-PSDB; AAQ98695.
 XX
 PT Yeast with reduced levels of aspartyl protease 3 proteolytic activity -
 PT used to secrete human albumin without prodn. of the 45 kD fragment.
 XX
 PS Example 1; Page 26-28; 50pp; English.
 XX
 CC The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected
 CC to site-directed mutagenesis to investigate the role of endoproteases in
 CC the generation of a 45 kDa albumin fragment obtd. when the cDNA is
 CC expressed in S. cerevisiae. Mutations were: R410A; L407A, L408V, V409A;
 CC and R410A, K413Q, K414Q. The latter set of mutations, especially,
 CC improved stability of HSA to yeast Yap3p proteolytic cleavage, allowing
 CC increased prodn. of recombinant HSA. (Updated on 25-MAR-2003 to correct
 CC PI field.)
 XX
 SQ Sequence 585 AA;
 Query Match 100.0%; Score 3103; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
 DB 1 DAHSEVAHRFKDLGEENFKALVLIAPFAQYLOQCPEFEDHVKLVNEVTEFAKTCVADESAAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVMCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 121 DVMCTAFHDNEETFLKKLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 QY 181 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVDTLTK 240
 DB 181 KLDELDRDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEVSCLVDTLTK 240
 QY 241 VHTECCHGDLLECADDDRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300
 DB 241 VHTECCHGDLLECADDDRADLAKYICENQDSISSKLKECEKPLEKSHCIAEVENDEMPA 300

QY 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
 DB 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVNCTAFHNDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180
 DB 121 DVNCTAFHNDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
 DB 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKYAEAKDVFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLEK 360
 DB 301 DLPSLAADFVESKDVCKYAEAKDVFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLEK 360
 QY 361 CAAADPHECYAKVDFEFKPLVEBPONLIKONCELFEQLGGEYFONALLVRYTKVPOVST 420
 DB 361 CAAADPHECYAKVDFEFKPLVEBPONLIKONCELFEQLGGEYFONALLVRYTKVPOVST 420
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540
 QY 541 KEQLKAVMDDFAAFVEKCKCKADKTCFAEKGKLVAAASQAALGL 585
 DB 541 KEQLKAVMDDFAAFVEKCKCKADKTCFAEKGKLVAAASQAALGL 585

RESULT 2

AA05318
 ID AAR05318 standard; protein; 585 AA.

AC AAR05318;
 DT 08-OCT-1990 (first entry)
 DE Human serum albumin gene product.
 KW Human serum albumin; HSA-A; yeast; ds.
 OS Homo sapiens.

PN JP02117384-A.
 PD 01-MAY-1990.
 PF 26-OCT-1988; 88JJP-00268302.
 PR 26-OCT-1988; 88JJP-00268302.
 PA (TOFU) TOA NENRYO KOGYO KK.
 XX WPI; 1990-176228/23.
 DR N-PSDB; AAQ04719.

PT Human serum albumin prepn. by yeast host - by culturing transformed
 PT plasmid yeast to produce serum, and removing it.
 PS Disclosure; Page ?; -pp; Japanese.

XX

CC Mature HSA-A may be produced using the sequence incorporated into a
 CC plasmid vector with suitable controllers, and transferred to a yeast
 CC expression system
 XX
 SQ Sequence 585 AA;
 Query Match 100.0%; Score 3103; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1.5e-254;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
 DB 1 DAHKEVAHRFKDLGSENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSEAE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
 QY 121 DVNCTAFHNDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180
 DB 121 DVNCTAFHNDNEETFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCOQADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
 DB 181 KLDELDEGKASSAKORLKCASLQKGERAFKAWAVARLSORFPKAEFAEVSCLVTDLT 240
 QY 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 241 VHTCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKYAEAKDVFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLEK 360
 DB 301 DLPSLAADFVESKDVCKYAEAKDVFLGMFLYEVARRHPDYSVLLLRLLAKTYETTLEK 360
 QY 361 CAAADPHECYAKVDFEFKPLVEBPONLIKONCELFEQLGGEYFONALLVRYTKVPOVST 420
 DB 361 CAAADPHECYAKVDFEFKPLVEBPONLIKONCELFEQLGGEYFONALLVRYTKVPOVST 420
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540
 DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQOTALVELVGHKPKAT 540
 QY 541 KEQLKAVMDDFAAFVEKCKCKADKTCFAEKGKLVAAASQAALGL 585
 DB 541 KEQLKAVMDDFAAFVEKCKCKADKTCFAEKGKLVAAASQAALGL 585

RESULT 3

AA08457
 ID AAR08457 standard; protein; 585 AA.

AC AAR08457;
 DT 25-MAR-2003 (revised)
 DT 16-APR-1991 (first entry)
 DE Human serum albumin.

XX HSA; folding; ss.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FT Region 1..303
 FT /label= B
 FT 123..585
 FT /label= C
 FT 123..303
 FT /label= A

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:24:29 ; Search time 344.086 Seconds
(without alignments)
480.375 Million cell updates/sec

Title: US-09-832-929-18
Perfect score: 3103
Sequence: 1 DAHKEVAHFKVLGHEHFK.....TCFAEKGKLIVAASQAALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_29Jan04:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	585	1 AAP90388	Mature hu
2	3103	100.0	585	2 AAR05318	Human ser
3	3103	100.0	585	2 AAR08457	Human ser
4	3103	100.0	585	2 AAR80301	Human ser
5	3103	100.0	585	2 AAO20111	HSA prote
6	3103	100.0	585	3 AAY84873	Amino aci
7	3103	100.0	585	3 AAY83946	Yeast cod
8	3103	100.0	585	4 AAM52567	Mature hu
9	3103	100.0	585	4 AAE12403	Human alb
10	3103	100.0	585	4 AAE13129	Human alb
11	3103	100.0	585	4 AAE13399	Human alb
12	3103	100.0	585	4 ABB79006	Human mat
13	3103	100.0	585	4 AAE08578	Human ser
14	3103	100.0	585	5 AAU75220	Mature fo
15	3103	100.0	585	5 ABJ00986	B lymphoc
16	3103	100.0	585	5 ABG33321	Human ser
17	3103	100.0	585	5 ABG33847	Human B, L
18	3103	100.0	585	5 ABG71291	Glycosyla
19	3103	100.0	585	6 ABR55695	Human alb
20	3103	100.0	585	7 ABR42606	Human ser
21	3103	100.0	585	7 ADC16767	Human ser
22	3103	100.0	585	7 ADD06469	Human ser
23	3103	100.0	585	7 ADD68016	Mature fo
24	3103	100.0	609	3 AAB36542	Recombina
25	3103	100.0	609	3 AAY78147	Pre human

26	3103	100.0	609	3 AAB36549	Recombina
27	3103	100.0	609	6 ABU57252	Human ser
28	3103	100.0	609	6 ABU57253	Human ser
29	3103	100.0	609	7 ADA24217	Human ser
30	3103	100.0	609	7 ADD06471	Human ser
31	3103	100.0	609	8 ADE77205	Human pro
32	3103	100.0	610	2 AAR39510	Chimeric
33	3103	100.0	616	6 AAE30916	Val8-GLP-
34	3103	100.0	624	6 AAE30919	Human ser
35	3103	100.0	631	6 AAE30917	Val8-GLP-
36	3103	100.0	640	6 AAE30920	Exendin-4
37	3103	100.0	640	6 AAE30918	Gly8-Glu2
38	3103	100.0	651	7 ADD06597	Human Ckb
39	3103	100.0	652	7 ADD06596	Human Ckb
40	3103	100.0	653	7 ADD06595	Human Ckb
41	3103	100.0	656	7 ADD06594	Human Ckb
42	3103	100.0	659	7 ADC16791	Chimeric
43	3103	100.0	659	7 ADC16793	Chimeric
44	3103	100.0	662	7 ADC16789	Chimeric
45	3103	100.0	662	7 ADC16787	Chimeric

ALIGNMENTS

RESULT 1
AAP90388
ID AAP90388 standard; protein; 585 AA.
XX
AC AAP90388;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Mature human serum albumin polypeptide.
XX
KW Human serum albumin; mature protein; new polypeptides; plasma expanders.
XX
OS Homo sapiens; (Human).
XX
PN EP322094-A.
XX
PD 28-JUN-1989.
XX
PF 25-OCT-1988; 88EP-00310000.
XX
PR 30-OCT-1987; 87GB-00025529.
XX
PA (DELZ) DELTA BIOTECHNOLOGY LTD.
XX
PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX
DR WPI; 1989-186464/26.
XX
DR N-ESDB; AAN90128.
XX
PT New N-terminal fragments of human serum albumin - esp. useful as blood plasma expanders.
XX
PS Disclosure; Fig 2; 20pp; English.
XX
CC Mature protein of human serum albumin (see corresp. AAN90128). Used to make new N-terminal fragments which are used as plasma expanders, or as substitutes for HSA or BSA, in tissue culture media. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)

Query Match 100.0%; Score 3103; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.5e-254;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_taxid=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;
RT "Partial mRNA sequence for tuatara A/B serum albumin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375973; AAM46106.1; -;
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR00264; Serum albumin.
DR Pfam; PF00273; transport prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PDC02486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN_2.
DR PROSITE; PS00212; ALBUMIN; 2.
FT NON TER 1
FT SEQUENCE 400 AA; 45715 MW; 8DE20609657CF753 CRC64;
SQ
Query Match 33.7%; Score 1045; DB 13; Length 400;
Best Local Similarity 47.7%; Pred. No. 2.8e-75;
Matches 186; Conservative 80; Mismatches 124; Indels 0; Gaps 0;
QY 196 QRLKASLQKFGERAFKAWARLSQRPKAEFAEVSKLVTDLTKVHTCCCHGLLECAD 255
DB 3 EKHSQGLKSGFERAFQANKLVLSQKFPKAPPEIHKVLTATKLVKQECCHGDMIECLD 62
QY 256 DRADLAKTICENQDSISSKLECCCKPLLEKSHCIAEVENDEMPADLPDLAADFVESKDV 315
DB 63 DRVENVAYICSKQAVFSSKIKCCCKPIVDREVCIQLQADLDEKPADLPISIAGQYIESTEV 122
QY 316 CKNYAEAKDVFLGMFLYVARHPDYVSVLLRLAKTYETTLKCCAAADPHECYAKVFD 375
DB 123 CKHYEKGKDVFLAHFVYYSRHPSPFSQMLRTGKGYQDILDKCKCTENPPECYKAGE 182
QY 376 EFKPLVERPQNLKQNCLEFQGLGEYKFNALLVRYTKKPOVSTPTLVEYSRNLGKVG 435
DB 183 ELARHQESQELIKTHCSFTSQGKDPFQWLVRYTKQFQLPABEIEISKLTGVGV 242
QY 436 KCKHPEAKRMPCAEDYLSVLNQLCVLHKTVPVSDRVTKCTESLVNRRPQFSALEVDE 495
DB 243 KCCFLSEDKRLSCSEKHLNWLFEICRQHEASPVNNHVTTCCTDSYSEMRPCFTKLGVD 302
QY 496 TVVPKEFNAETTFHADICTLSEKROIKQATLVKGVKPKATKQOLKAVMDDFAAV 555
DB 303 SYVPEFCPSTFLDFQLCTAPEEARLKKQLTLVKKLIQLKQIEDQKLVTDYHAME 362
QY 556 EKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 363 EKCCQAKENKQECFSTEGEKLQEGKALLGV 392

Search completed: April 19, 2004, 11:59:55
Job time : 228.789 secs

DR GO:0001542; P:ovulation (sensu Mammalia); IMP.
 DR GO:0042448; P:progestrone metabolism; IMP.
 DR GO:0019953; P:sexual reproduction; IMP.
 DR InterPro: IPR000264; Serum albumin.
 DR Pfam: PF00273; transport prot; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 2.
 SQ SEQUENCE 605 AA; 67322 MW; 048B7A4A8B01E44B CRC64;

Query Match 34.9%; Score 1082; DB 11; Length 605;
 Best Local Similarity 35.7%; Pred. No. 5.1e-78;
 Matches 204; Conservative 120; Mismatches 241; Indels 6; Gaps 3;

QY 16 EENFKALVLIAPAYLQOCPEFHVKNVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 75
 DB 40 EKNVLSIATITFTQFPEATEEE----VNKNTSDVLAAMKNGSGGCLSESQSVFDEIC 95

QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDNP-NLPRLVPEVDVNMCTAFPHNEETFF 134
 DB 96 HETELSNKYG-LSGCCSQSGVERHQCLLARKKTAPASVPPFPPEPABSCKAHEENRAVF 154

QY 135 LKXYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLPKLDELDEGKASSA 194
 DB 155 MNRFIYVSRNRPWYAPAILSLAAQYDKVYLACCKADNKECFQTKRASIAREGSM 214

QY 195 KORLKCASLOKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTQVHTECHGDLLECA 254
 DB 215 LNEHVCSVIRKFSRNLQATTIIKLSQKLTANFTETQKALDVALIHEECCQGSLECL 274

QY 255 DDRADLAKYICENODSISLKECCBPKLEKSHCAEVENDEMPADLPSLAADFVESKD 314
 DB 275 QDEKVMYIYCSQNISSKIAECCKPMQLGFCIIHAENGVPKGLSINPSQFLGDRN 334

QY 315 VCKNYAARADVFLGMFLYBYARRHPDYVSVLLRLAKYTTETLTKCAAADPHECYAKVF 374
 DB 335 PAQFSSEKIMFMASFLHEYSRTHPNLPVSVILRIAKTYQEIILEKCSQSGNLPQCQDNLE 394

QY 375 DEKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKVPQVSTPTLVEVSRNLKVG 434
 DB 515 EYAPPPFSEDKFIFHKDLCOAHKALQTKWQELLINLVKQKPELIEQLAAVTADFSGL 574

QY 555 VEKCKADDKETCFPAEKGKLVAAASQAALGL 585
 DB 575 LEKCKAQDQEVCFTEEGPKLISKTRDALGV 605

RESULT 14
 Q8BK56 PRELIMINARY; PRT; 605 AA.
 ID Q8BK56
 AC Q8BK56;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Alpha-fetoprotein precursor.
 GN AFP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
 RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 DR EMBL; AK076197; BAC36249.1; -.
 DR MCD; MGI:87851; Afp.
 DR GO:0005815; C:extracellular space; IDA.
 DR GO:0001542; P:ovulation (sensu Mammalia); IMP.
 DR GO:0042448; P:progestrone metabolism; IMP.
 DR GO:0019953; P:sexual reproduction; IMP.
 DR InterPro: IPR000264; Serum albumin.
 DR Pfam: PF00273; transport prot; 3.
 DR PRINTS: PR00802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum albumin; 1.
 DR SMART: SM00103; ALBUMIN; 3.
 DR PROSITE: PS00212; ALBUMIN; 2.
 SQ SEQUENCE 605 AA; 67409 MW; CF3509A9EC146110 CRC64;

Query Match 34.7%; Score 1076; DB 11; Length 605;
 Best Local Similarity 35.6%; Pred. No. 1.5e-77;
 Matches 203; Conservative 120; Mismatches 242; Indels 6; Gaps 3;

QY 16 EENFKALVLIAPAYLQOCPEFHVKNVNEVTEFAKTCVADESAENCDKSLHTLFGDKLC 75
 DB 40 EKNVLSIATITFTQFPEATEEE----VNKNTSDVLAAMKNGSGGCLSESQSVFDEIC 95

QY 76 TVATLRETYGEMADCCAKQEPERNECFLOHKDNP-NLPRLVPEVDVNMCTAFHNEETFF 134
 DB 96 HETELSNKYG-LSGCCSQSGVERHQCLLARKKTAPASVPPFPPEPABSCKAHEENRAVF 154

QY 135 LKXYLVEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLPKLDELDEGKASSA 194
 DB 155 MNRFIYVSRNRPWYAPAILSLAAQYDKVYLACCKADNKECFQTKRASIAREGSM 214

QY 195 KORLKCASLOKQGERAFKAWAVARLSQRPFAEFAVSKLVTDLTQVHTECHGDLLECA 254
 DB 215 LNEHVCSVIRKFSRNLQATTIIKLSQKLTANFTETQKALDVALIHEECCQGSLECL 274

QY 255 DDRADLAKYICENODSISLKECCBPKLEKSHCAEVENDEMPADLPSLAADFVESKD 314
 DB 275 QDEKVMYIYCSQNISSKIAECCKPMQLGFCIIHAENGVPKGLSINPSQFLGDRN 334

QY 315 VCKNYAARADVFLGMFLYBYARRHPDYVSVLLRLAKYTTETLTKCAAADPHECYAKVF 374
 DB 335 PAQFSSEKIMFMASFLHEYSRTHPNLPVSVILRIAKTYQEIILEKCSQSGNLPQCQDNLE 394

QY 375 DEKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKVPQVSTPTLVEVSRNLKVG 434
 DB 395 EEOKHEESQALSKSCALYQTLGDKLQNLFLGYTKAPQLTSAELIDITGKMWSTA 454

QY 435 SKCKHPEAKMPCAEYLSVNLQNLVHKEKTPVSDRVTKCTESLVNRRPFCFSALEVD 494
 DB 455 STCCQLSEKWSGGEGMADIFIGHLCIRNEASPVNSGISHCCNSYSNRLCITSFLRD 514

QY 495 EYVPEFNAETFTFHADICTLSEKEROIKKOTALVELVKHKPKATKQKAVMDPFAAF 554
 DB 515 EYAPPPFSEDKFIFHKDLCOAHKALQTKWQELLINLVKQKPELIEQLAAVTADFSGL 574

QY 555 VEKCKADDKETCFPAEKGKLVAAASQAALGL 585
 DB 575 LEKCKAQDQEVCFTEEGPKLISKTRDALGV 605

RESULT 15
 Q8JIA7 PRELIMINARY; PRT; 400 AA.
 ID Q8JIA7
 AC Q8JIA7;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE A/B over-sized serum albumin (Fragment).
 OS Sphenodon punctatus (Hatteria) (Tuatara).

```

QY 16 EENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSAENCDKSLHTLFGDKLC 75
Db 40 EMNLVDLATIFFAQVHGATYETIRKMWANDVLAVIKPTDGEHPTGCLNQMSAFLEEC 99
QY 76 TVATLRTEYEMADCCAKQPERNECFLOHKDNDPNLRLVRPEVDVMTAFHDNEETFL 135
Db 100 HEKIFIKYQ-LSOCCRSSEERSCLPAHKAAPALPPQAPPEPATSCKAYEEDRETFM 158
QY 136 KKYLYEIAARRHPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPY 195
Db 159 NKCIYEISRRHPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPY 218
QY 196 ORLKASLOKGERAFKAWARLSQRPFAEFAEVSQKLVDTLTKVHTECHGDLLECAD 255
Db 219 NOHCAVKGRTPEELIFQAITVAKLSQRPFRANFTBIOKLVDVAHIHEECRGNVLECLQ 278
QY 256 DRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVNDEMPPADLPSLAADPVEKOV 315
Db 279 DGEKIMSYICSGOALSCKIAECCCKPPLLEKSHCIAEVNDEMPPADLPSLAADPVEKOV 338
QY 316 CKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTLKCCAAADPHECYAKVFD 375
Db 339 NOFSAEENIFLASVHYSRRHPDLAVPILRVKGVQEVLEKCAOSEKPLECQDKEE 398
QY 376 EFKPLVEEPQNLKONCELFQGEYKFNALLVRYTKVPQVSTPTLVEVSRNLGKVG 435
Db 399 ELQKTSQSOAMAKSCGLFQKLGQYVLQNAFLVSVYKVPOLTSIAELMALTEKMAATAA 458
QY 436 KCKHPEAKRMPCAEDYLSVNLQNLVHEKTPYSDRYTKCTESLVNRRPCSALEVDE 495
Db 458 TCCQLSEKQLACGGAADLIIGHLCIRHEVNPVNPVGVHCNCSYANRRPCFSLSLIDE 518
QY 496 TVVPKEFNAETFTFHADICTLSEKROIKKQATLVKPKATKEQLKAVMDPFAFV 555
Db 519 TVVPSFSADKIFPKDLCOPOGVALQTMQBFNLVYKQFQITEQLEAVIADFSGLL 578
QY 556 EKCKXADDKETCFABEGKLVKVAASQALGL 585
Db 579 ETCCQEQEQEVCFABEGPKLISKTRDALGV 608

RESULT 12
Q8UW06 PRELIMINARY; PRT; 624 AA.
ID Q8UW06;
AC Q8UW06;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ambystoma texanum (Smallmouth salamander).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
OC Ambystoma.
OC NCBI_TaxID=8304;
RN [1]
RP SEQUENCE FROM N.A.
RA Haverfield B.V., Uzzell T., Spolsky C.M., Bazartseren B.;
RT "Serum albumin of the mole salamanders Ambystoma maculatum and
RT Ambystoma texanum."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217182; AAL56645.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 1.
DR KEGG; K01101; ALBUMIN.
DR SIGNAL.
FT SIGNAL 1 24 POTENTIAL.

```

```

FT CHAIN 25 624 SERUM ALBUMIN.
SQ SEQUENCE 624 AA; 70321 MW; DE08533BF4953EF7 CRC64;

Query Match 35.0%; Score 1087; DB 13; Length 624;
Best Local Similarity 37.7%; Pred.No. 2.1e-78; Indels 12; Gaps 5;
Matches 214; Conservative 102; Mismatches 240;

QY 14 LGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADSAENCDKSLHTLFGDK 73
Db 45 IGVEHAKALAMALFQMLSKCPHHEQVQVRVNMVMDIADLCGRGAKHGDCGKSVMTIILNE 104
QY 74 LCTVATLRTEYEMADCCAKQPERNECFLOHKDNDP-NLRLVRPEVDVMTAFHDNEE 132
Db 105 ICKTPENPEKYPFHEGCKCKEDPERHKCFIEHKSDDPKERTYEVKPSPEQICKOHAENRD 164
QY 133 TFLKLYLXIARRHPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPYAPY 191
Db 165 EFLGHYIHKVASHHTWYPPAILSTLHPDGVSHCCCKDEATVGQCLSEKMPAHKEVEH 224
QY 192 SSAKQRLKASLOKGERAFKAWARLSQRPFAEFAEVSQKLVDTLTKVHTECHGDL 251
Db 225 VCAVQKHNCYIQLQNFNERALRASKAHACSKPFPASFENVQRLTDGIVHLHQTCCGGNM 284
QY 252 ECADDRADLAKYICENQDSISSKLECCCKPPLLEKSHCIAEVNDEMPPADLPSLAADPVE 311
Db 285 ACMAERKLUPTQCEK-----KCCCKPVLSECIIVRLPNDKPADLSPVRYFYD 336
QY 312 SKDVCKNYAEAKDVFLGMFLYIYARRHPDYSVLLRLAKTYETTLKCCAAADPHECYA 371
Db 337 DPEVCKRFKEEGDAFWGRFLCDYAKIHPHSAELNLRISAGLEKAYKTCCEGAHNECIA 396
QY 372 KVDFEPKPLVEEPQNLKONCELFQGEYKFNALLVRYTKVPQVSTPTLVEVSRNLG 431
Db 397 KEBETLRHEIASEKTKLTKTCGALEKLGYPHFONIMIVRYTGLPRSSDAFLIYIKTLT 456
QY 432 KVGSKCKCHPEAKRMPCAEDYLSVNLQNLVHEKTPV-SDRYTKCTESLVNRRPCFSA 490
Db 457 NIGOKCKLPEDQMPCEGGLGWVFAQIC-QNOKTPFENEKLAHCKDSLSFTTCFAA 515
QY 491 LEVDETVVPKEFNAETFTFHADICTLSEKROIKKQATLVKPKATKEQLKAVMD 550
Db 516 LTVDETVVPKAVTAESFNFNDEFTSEADLQAKOTFLHLVTRHPKITDEQVTKISEK 575
QY 551 FAFVEKCKKADDKETCFABEGKLVKVAAS 578
Db 576 FLAMCGCKCKAQDQNECFATEGAKLVEA 603

RESULT 13
Q8BK65 PRELIMINARY; PRT; 605 AA.
ID Q8BK65;
AC Q8BK65;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Alpha-fetoprotein precursor.
GN AFP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK076053; BAC36150.1; -.
DR MGD; MGI:87951; Afp.
DR GO; GO:0005615; C:extracellular space; IDA.

```



```

DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 417 AA; 47360 MW; 1687648333EFAE8D CRC64;

Query Match 64.2%; Score 1991; DB 4; Length 417;
Best Local Similarity 95.2%; Pred. No. 6.5e-151;
Matches 381; Conservative 3; Mismatches 10; Indels 6; Gaps 1;

QY 186 RDEGKASSAKORLKASLOKFGERAFAKAWARLSORPPKAEFAEVS KLVDLT LTKVHTEC 245
DB 24 RDAKSEVAHR-----FKDLGEENFKAWARLSORPPKAEFAEVS KLVDLT LTKVHTEC 77
QY 246 CHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADIPSL 305
DB 78 CHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPADIPSL 137
QY 306 AADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRAKTYETTTLEKCAAAD 365
DB 138 AADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRAKTYETTTLEKCAAAD 197
QY 366 PHECYAKVDFEKPVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVSTPTLVE 425
DB 198 PHECYAKVDFEKPVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVSTPTLVE 257
QY 426 VSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTESLVNRR 485
DB 258 VSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTESLVNRR 317
QY 486 PCFSALEVDVETVPKFEFNAETFTFHADICTLSEKSRQIKKQATALVELVKHKPKATKEOLK 545
DB 318 PCFSALEVDVETVPKFEFNAETFTFHADICTLSEKSRQIKKQATALVELVKHKPKATKEOLK 377
QY 546 AVMDDFAAFAVEKCKCKADDETCFAEKGKLVAAASQAALGL 585
DB 378 AVMDDFAAFAVEKCKCKADDETCFAEKGKLVAAASQAALGL 417

RESULT 6
Q81UK7 PRELIMINARY; PRT; 396 AA.
AC Q81UK7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to serum albumin precursor.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035969; AAH35969.1; -
DR GO; GO:0005815; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

Query Match 60.1%; Score 1865.5; DB 4; Length 396;
Best Local Similarity 63.6%; Pred. No. 6.7e-141;
Matches 372; Conservative 0; Mismatches 0; Indels 213; Gaps 1;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 DAHKSEVAHRFKDLGEENFKALVLIAPQYLOQCPEFDHVKLVNEVTEFAKTCVADESAAE 84

us-09-832-929-18.ispt
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHNDNEETFLKKLYEYARRHPYFYAPELLFFAKRYKAAFTCCQCAADKAAALP 180
DB 145 DVNCTAFHNDNEETFLKKYL----- 163
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFAKAWARLSORPPKAEFAEVS KLVDLT LTK 240
DB 164 ----- 163
QY 241 VHTCECHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 164 ----- 163
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRAKTYETTTLEK 360
DB 164 -----YETTTLEK 171
QY 361 CAADPHCEYAKVDFEKPVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 172 CAADPHCEYAKVDFEKPVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKVPQVST 231
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
DB 232 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCTES 291
QY 481 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKSRQIKKQATALVELVKHKPKAT 540
DB 292 LVNRRPCFSALEVDVETVPKFEFNAETFTFHADICTLSEKSRQIKKQATALVELVKHKPKAT 351
QY 541 KEQLKAVMDDFAAFAVEKCKCKADDETCFAEKGKLVAAASQAALGL 585
DB 352 KEQLKAVMDDFAAFAVEKCKCKADDETCFAEKGKLVAAASQAALGL 396

RESULT 7
Q8JIA9 PRELIMINARY; PRT; 527 AA.
AC Q8JIA9
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Serum albumin (Fragment)
OS Sphenodon punctatus (Tuatara)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
[1]
RN SEQUENCE FROM N.A.
RC Metcalf V.J.; Brennan S.O.; George P.M.; Chambers G.K.;
RA "Partial coding sequence for Sphenodon punctatus 68 kDa albumin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375971; AA046104.1; -
DR GO; GO:0005815; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON TER
SQ SEQUENCE 527 AA; 59711 MW; C62B799B387F5929 CRC64;

Query Match 41.7%; Score 1295.5; DB 13; Length 527;
Best Local Similarity 45.9%; Pred. No. 3.6e-95;
Matches 240; Conservative 92; Mismatches 188; Indels 3; Gaps 3;

QY 62 CDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

```

```

RESULT 4
Q8C7C7
ID ID PRELIMINARY; PRT; 576 AA.
AC AC
Q8C7C7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin 1 (Fragment).
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR ENBL; AK050644; BAC34360.1; -.
DR DR
DR MGI; 87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0003386; F:carrier activity; IEA.
DR GO; GO:0008100; P:transport; IEA.
DR InterPro; IPR000284; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.

```

DR PROSITE; PS00212; ALBUMIN; 3.
FT NON_TER 1
SQ SEQUENCE 576 AA; 65002 MW; F85733E99AE37F04 CRC64;

Query Match 75.1%; Score 2336; DB 11; Length 576;
Best Local Similarity 72.2%; Pred.No. 2.5e-178;
Matches 415; Conservative 79; Mismatches 81; Indels 0; Gaps 0;

Qy 9 HRFKDLGEENFKALVLIAFAYLQQCPFDHVKLVNEVTEFAKTCTVADESAENCDSLSHT 68
:
Dd 1 NRYNDLGEQHFGKLVLIAFSOYLQKCSYDBAKLVQEVDTFDKTVADESAANCDSLSHT 60
:
Qy 69 LFGDKLTATNLETGCEWADCCAKOBPENECFLOHKDDNPMLPLRVREVDVMCTAFH 128
:
Dd 61 LFGDKLCALPNLRNTYGELADCTCKOBPENECFLOHKDDNPSLPPERPEEAEMACTSFK 120
:
Qy 129 DNBETFLKKLYIARIHRHPFYFAPELLFFAKRKYKAFTCCQAADKAACLLPKLDLRDE 188
:
Dd 121 ENPTFMGHYLEHVARRHPFYFAPELLYYAEQYNEILTQCCEADKESCLTPKLDGVKEK 188
:
Qy 189 GKASSAQRLKCASLOKFGERAKANAVARLSORFPKAEFAEVSKLVDTIKVHTECHG 248
:
Dd 181 ALVSSVRQRMKCSSMQKFGERAFKAWARLSOTFFNADFAETIKLATDILTKNKECCHG 240
:
Qy 249 DLLECADDRADLAKEYICENDODSISSLKKECCXPLEKSHCTAENVENDMPADLPISAAD 308
:
Dd 241 DLLECADDRADLAKEYICENOATISSLKQTCDFAPLLKKAHLSEVEHDTPADLPATAAD 300
:
Qy 309 FVSQDVCKNYBAKDVFGLGWFIYEYARRHPDYSVVLLLRSLAKTYETTLKCCAAADPHE 368
:
Dd 301 FVEDQECVKNYAERAKDVFGLFTFLEYERRHPDYSVSLLLRLAKYEAATLEKCCAANPPA 360
:
Qy 369 CYAKVPDEFKPLVEEPONLIKQNCCLPEQLGEYKFNALLVRYTKYPVOYSTPTLVEVSR 428
:
Dd 361 CYTGLVAEFPLVEEPKNLVKTNCDLYEKUGGYGNAILIVRYTKAPQVSTPTLVEEAR 420
:
Qy 429 NLGKVSGKCKCHPEAKMPCAEDYSVLNLQCLVLEHKTVPDSRVTKCTESTLYNNRPFCF 488
:
Dd 421 NLGRVGTKCTLPEDQRLPCVEDYLSAILNRVCLLHEKTPVSEHVTKCCSGSLVERRPCF 480
:
Qy 489 SALEVDETYPKPFNAETPTFHADICTLSKERQIKKOTALVELYKHKPAATKEOLKAVM 548
:
Dd 481 SALTVDETYYPKPFKAETPTFHSDICTLPEKEQIKKOTALABELVHKHPKATABQLKTVM 540
:
Qy 549 DDFAAFVEKCKCKADDKETCFABEGKKLVAASAAL 583
:
Dd 541 DDFAQFLDTCCKAADKDTCTESTGPNLVTECKDAL 575
: :

RESULT 5
Q86YGO PRELIMINARY; PRT; 417 AA.

AC Q86YGO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to alpha-fetoprotein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041789; AAH41789.1; -
DR GO; GO:0005615; Cytoplasmic space; IEA.
DR GO; GO:0005386; F-carrier activity; IEA.
DR GO; GO:0006810; Protein transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00402; SERUMALBMIN

```
121 DAVCTAFHNEEFTLKKLYIEIARRHPYFAPPELLYAEYRGVTECEADKAACLTIP 180
181 KLDELDEGKASSAKORLKCASLOKFGGRAPKAWAVARLSORFPAEYAKVSKLVTDLT 240
181 KYDALREKVLASSAKERLKCASLOKFGGRAPKAWAVARLSORFPAEYAKVSKLVTDLT 240
241 VHTCHGDLLECCADRADLAKYICENODSISSKLEKCEKPLLEKSHCIAFVENDMPA 300
241 IHKECHGDLLECCADRADLAKYICENODSISSKLEKCEKPLLEKSHCIAFVENDMPA 300
301 DPLSLAADFVSKDVKCYAEAKDVLGMLFVEYARRHPDYSVLLLRKATYETLEK 360
301 DPLSLAADFVSKDVKCYAEAKDVLGMLFVEYARRHPDYSVLLLRKATYETLEK 360
361 CAADPHCEYAKVDFEKLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
361 CATDDPACAYAFVDFEKLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEADYLSVNLQCVLHKTVPVSDRVTKCCTES 480
421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEADYLSVNLQCVLHKTVPVSDRVTKCCTES 480
481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVHKHPKAT 540
481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVHKHPKAT 540
541 KEQKAVMDPFAFVEKCCADKCTCPABEGKLVAAQAAL 583
541 EBQLTKVMGDFGSDVFDKCAAREKACFABEGKLVAAQAAL 583
```

RESULT 2

```
Q95VB7
ID Q95VB7 PRELIMINARY; PRT; 508 AA.
AC Q95VB7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Staderker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RT component in murine infection with Schistosoma mansoni";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418550; AL008579.1; -.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
SQ SEQUENCE 608 AA; 68225 MW; E5EABB28E1C66B54 CRC64;
```

Query Match 80.7%; Score 2504; DB 5; Length 608;
Best Local Similarity 76.3%; Pred. No. 9.9e-192;
Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

```
1 DAHSEVAHRFKDLGBENFKALVLIAPQYLOQCPEFDHVKLVNEVTEFAKTCVADESAA 60
25 DAHSEVAHRFKDLGBENFKALVLIAPQYLOQCPEFDHVKLVNEVTEFAKTCVADESAA 84
61 NCDKSLHTLFGDKLCTVATREYVGMADCCAKQEPERNECTLOHDDNPNLPRLVREPV 120
85 NCDKSLHTLFGDKLCTVATREYVGMADCCAKQEPERNECTLOHDDNPNLPRLVREPV 144
```

Query Match 76.5%; Score 2374; DB 11; Length 608;
Best Local Similarity 72.2%; Pred. No. 2.5e-181;
Matches 421; Conservative 80; Mismatches 82; Indels 0; Gaps 0;

```
121 DVNCTAFHNEEFTLKKLYIEIARRHPYFAPPELLYAEYRGVTECEADKAACLTIP 180
145 EAMCTSGFQENAVTFMGHYLHEVARRHPYFAPPELLYAEYRGVTECEADKAACITP 204
181 KLDELDEGKASSAKORLKCASLOKFGGRAPKAWAVARLSORFPAEYAKVSKLVTDLT 240
205 KLDELDEGKASSAKORLKCASLOKFGGRAPKAWAVARLSORFPAEYAKVSKLVTDLT 240
241 VHTCHGDLLECCADRADLAKYICENODSISSKLEKCEKPLLEKSHCIAFVENDMPA 300
265 LTBECCHGDLLECCADRADLAKYICENODSISSKLEKCEKPLLEKSHCIAFVENDMPA 324
301 DPLSLAADFVSKDVKCYAEAKDVLGMLFVEYARRHPDYSVLLLRKATYETLEK 360
325 DPLSLAADFVSKDVKCYAEAKDVLGMLFVEYARRHPDYSVLLLRKATYETLEK 384
361 CAADPHCEYAKVDFEKLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 420
385 CAADPHCEYAKVDFEKLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKVPQVST 444
421 PTLVEVRNLGKVGSKCCCHPEAKMPCAEADYLSVNLQCVLHKTVPVSDRVTKCCTES 480
445 PTLVEVRNLGKVGSKCCCHPEAKMPCAEADYLSVNLQCVLHKTVPVSDRVTKCCTES 504
481 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVHKHPKAT 540
505 LVNRRPCFSALEVDETYVPKFEFNAETFTFHADICTLSEKERQIKKOTALVELVHKHPKAT 564
541 KEQKAVMDPFAFVEKCCADKCTCPABEGKLVAAQAAL 583
565 GPQLRTVLGFEFTFLDKCKCAEDKACFSDGSPKLVASSQAAL 607
```

RESULT 3

```
Q8C7H3
ID Q8C7H3 PRELIMINARY; PRT; 608 AA.
AC Q8C7H3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Albumin 1.
GN ALB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Liver;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK050248; BAC34145.1; -.
DR MGI; MGI:87991; Alb1.
DR GO; GO:0005615; C:extracellular space; IEA.
DR GO; GO:0005386; F:carrier activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot.; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
SQ SEQUENCE 608 AA; 68722 MW; 292FE00BED3A61B4 CRC64;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:09 ; Search time 225.789 Seconds
(without alignments)
817.479 Million cell updates/sec

Title: US-09-832-929-18
Perfect score: 3103
Sequence: 1 DAHSEVAHRFKDLGEENFK.....TCPAEGKLVLAASQAALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2614	84.2	584	Q7YSG3	Q7YSG3 felis silve
2	2504	80.7	608	Q95VB7	Q95VB7 schistosoma
3	2374	76.5	608	Q8C7H3	Q8C7H3 mus musculus
4	2336	75.3	576	Q8C7C7	Q8C7C7 mus musculus
5	1991	64.2	417	Q86YGO	Q86YGO homo sapien
6	1865.5	60.1	396	Q81UK7	Q81UK7 homo sapien
7	1295.5	41.7	527	Q81IA9	Q81IA9 sphenodon p
8	1242	40.0	609	Q8MJU5	Q8MJU5 canis famil
9	1242	40.0	626	Q8UW05	Q8UW05 ambystoma m
10	1218.5	39.3	610	Q8MJ76	Q8MJ76 sus scrofa
11	1183.5	38.1	608	Q7TSF3	Q7TSF3 marmota mon
12	1087	35.0	624	Q8UW06	Q8UW06 ambystoma t
13	1082	34.9	605	Q8BK65	Q8BK65 mus musculus
14	1076	34.7	605	Q8BK56	Q8BK56 mus musculus
15	1045	33.7	400	Q81IA7	Q81IA7 sphenodon p
16	955	30.8	603	Q9YGH6	Q9YGH6 rana shqipe

17	928.5	29.9	614	13	Q91134	Q91134 naja naja
18	888	28.6	406	13	Q8JIA8	Q8JIA8 hoplodactyl
19	759	23.8	205	11	Q8CG74	Q8CG74 mus musculus
20	713	23.0	417	11	Q8R0J9	Q8R0J9 mus musculus
21	376.5	12.1	484	13	Q9W6FS	Q9W6FS gallus gall
22	372	12.0	491	11	Q7TS97	Q7TS97 mus musculus
23	365	11.8	551	13	Q42279	Q42279 petromyzon
24	331	10.7	122	13	Q90WZ8	Q90WZ8 larus argen
25	290	9.3	123	13	Q90WZ6	Q90WZ6 poephila gu
26	284	8.5	135	11	Q63205	Q63205 rattus norv
27	196	6.3	807	2	Q83ZS7	Q83ZS7 helicobacte
28	188	6.1	1723	2	Q9JMX8	Q9JMX8 helicobacte
29	184	5.9	1819	16	Q9ZLVO	Q9ZLVO helicobacte
30	184	5.9	1927	16	Q25262	Q25262 helicobacte
31	162.5	5.2	44	6	Q95WC2	Q95WC2 equus cabal
32	162.5	5.2	680	5	Q9V6S8	Q9V6S8 drosophila
33	161	5.2	1079	3	Q94V11	Q94V11 pneumocysti
34	156	5.0	1026	3	Q74569	Q74569 pneumocysti
35	154.5	5.0	3843	5	Q9USD0	Q9USD0 drosophila
36	153.5	4.9	661	5	Q8MS79	Q8MS79 drosophila
37	153.5	4.9	3843	5	Q9VU94	Q9VU94 drosophila
38	153	4.9	62	6	Q8MIL1	Q8MIL1 sus scrofa
39	150	4.8	1065	3	Q01828	Q01828 pneumocysti
40	149	4.8	40	6	Q9TRA5	Q9TRA5 oryctolagus
41	148	4.8	1028	3	Q74568	Q74568 pneumocysti
42	142.5	4.6	2756	10	Q9LJ60	Q9LJ60 arabidopsis
43	141.5	4.6	1560	5	Q26644	Q26644 strongyloce
44	138.5	4.5	1069	3	Q96V12	Q96V12 pneumocysti
45	137.5	4.4	1348	16	Q8YK55	Q8YK55 anabaena sp

ALIGNMENTS

RESULT 1

Q7YSG3 PRELIMINARY; PRT; 584 AA.
 ID Q7YSG3
 AC Q7YSG3
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Albumin (Fragment).
 GN ALB.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Reininger R., Swoboda I., Bohle B., Hauswirth A.W., Valent P.,
 RA Rumpold H., Valenta R., Spitzauer S.,
 RT "Escherichia coli expression and purification of recombinant cat
 albumin: IgE recognition, induction of basophil activation and
 lymphoproliferative responses in atopic patients."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ487677; CAD32275.1; -
 FT NON TER
 FT SEQUENCE 584 AA; 65908 MW; B51002F12902C9CE CRC64;

Query Match 84.2%; Score 2614; DB 6; Length 584;
 Best Local Similarity 81.6%; Pred. No. 1.5e-200;
 Matches 476; Conservative 54; Mismatches 53; Indels 0; Gaps 0;

Qy	1	DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPDHVKLVNEVTEFAKTCVADESAAE 60
Db	1	EAHGEIAHRFNDLGEHFGLVLAFAVSQVLOCCPDHVKLVNEVTEFANGCVADOSAA 60
Qy	61	NCDKSLHTLFGDKLCTVATLRETYGENADCCAKOBERNECFLOHKDDNNLRLVPEV 120
Db	61	NCDKSLHTLFGDKLCTVATLRETYGENADCCAKOBERNECFLOHKDDNNLRLVPEV 120
Qy	121	DMVCTAFDNEETLKKVLYEIARRHFFYFAPELLFFAKRYKAAFTCCCAADKAACLLP 180

```
FT DISULFID 302 313
FT DISULFID 384 393
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
FT CARBOHYD 251 251
FT
FT N-LINKED (GLCNAC. . .).
FT /FTID=CAR_000070.
FT A -> G (in dbSNP:1057173).
FT /FTID=VAR_012049.
SQ SEQUENCE 609 AA, 68677 MW, 4D9E45820E1C2D4F CRC64;

Query Match 40.3%; Score 1249.5; DB 1; Length 609;
Best Local Similarity 39.9%; Pred. NO. 5.5e-73;
Matches 235; Conservative 116; Mismatches 231; Indels 7; Gaps 3;

QY 3 HXSE-----VAHFRDGLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTBFAKTCVADE 57
DB 22 HREYGLASILDSYQCTAEISLADLTIFFAQFVEATYKEVSKVAKDALTAEKPTGDE 81
QY 58 SAENCDKSLHTLFGKLCVTATRETYGEMADCCAKQEPERNECPLOHKDNP-NLPLV 116
DB 82 QSSGCIENQLPAFLBELCHEKEILEKYGH-SDCCSQSBEGRHNCPLAHKXPTPASIPLFQ 140
QY 117 REVDVMCTAFHNDRETFKXLYLRIARHPYFYAPELLFPKRYKAAFTBCCQAADKAA 176
DB 141 VPEPVTSCAYBEDRETFNKFYIARHHPFLYAPTILLWAARYDKIIPSCCAENAVE 200
QY 177 CULPKLEURDEGKASSAQRKLCASLQKGERAFKAWARLSORFPKAEFAEVSCLVT 236
DB 201 CQTAKAATVTKELRESSLLNQHACAVMKNFGTRTQAIVTKLSQKTKVNFTEIQKLV 260
QY 237 DLTKVTECHGDLLECADDRADLAKYICENQDSISSKKECCEKPLEKSHCIAEVEND 296
DB 261 DVAHVHEHCRCGVDLCLQDGKIMSYICSQDITLSNKITECKLTLLERGQCIIHAEND 320
QY 297 EMPADLPSLAADPVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYVSVLLLLAKTYET 356
DB 321 EKPEGLSPNINRFLGDRDNQFSSGGEKNIFLASFVHEYSRHPQLAVSVILRAVKGQEL 380
QY 357 LEKCCAAADPHECYAKVDFEPLVEEPQNLKQNCLEPQELGEYKFNALLVRYTKVP 416
DB 381 LEKCFOTENPLECQDGEBELQYIOESQALAKRSQGLFQKLGYYLQNAFLVAYTKKAP 440
QY 417 QVSTPTLVEYSRNLGVSKCKKHPKAEADYLSVVLNQLCVLHEKTPVSDRVTKC 476
DB 441 QLTSSSELMAITRKGAATAATCCQLSEDKLLACGGAADIIGHLCIRHEWTPNFGVGQC 500
QY 477 CTESLVNRRPCFSGALEYDITYPKFNAETFTFHADICTLSEKEROIKQTALVELVKHK 536
DB 501 CTSSYANRRPCFSSLVVDITYVPPAFSDDKFIFHKDLQQAQGVALQTMKQEFILNLVKQK 560
QY 537 PKATKEQLKAYMDDFAAFVEKCKKADDKETCFAEKGKLVAAASQAALGL 585
DB 561 PQITEQLEAVIADFGLEKCKCQGGQEQEVCFABEGQKLSIKTRAAALGV 609
```

Search completed: April 19, 2004, 11:52:43
Job time : 44.6731 secs

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
FAhey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalilus D.B.,
RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=93278385; PubMed=7684942;
RA McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M.,
RA Tilghman S., Krumlauf R., Tuddenham E.G.D.;
RT "A G-->A substitution in an HNF I binding site in the human alpha-
fetoprotein gene is associated with hereditary persistence of alpha-
fetoprotein (HPAFP).";
RL Hum. Mol. Genet. 2:379-379(1993).
RN [5]
RP SEQUENCE OF 429-556 FROM N.A.
RX MEDLINE=91242409; PubMed=1709810;
RA Pucci P., Siciliano R., Malorni A., Marino G., Tecce M.F.,
RA Caccarini C., Terrana B.;
RT "Human alpha-fetoprotein primary structure: a mass spectrometric
study.";
RL Biochemistry 30:5061-5066(1991).
RN [7]
RP PRELIMINARY SEQUENCE OF 19-35.
RX MEDLINE=77242506; PubMed=70228;
RA Yachnin S., Heu R., Heinrikson R.L., Miller J.B.;
RT "Studies on human alpha-fetoprotein. Isolation and characterization
of monomeric and polymeric forms and amino-terminal sequence
analysis.";
RL Biochim. Biophys. Acta 493:418-428(1977).
RN [8]
RP PRELIMINARY SEQUENCE OF 19-38.
RX MEDLINE=78001760; PubMed=71198;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Comparative chemical structures of human alpha-fetoproteins from
fetal serum and from ascites fluid of a patient with hepatoma.";
RL Cancer Res. 37:3663-3667(1977).
RN [9]
RP PRELIMINARY SEQUENCE OF 19-39.
RX MEDLINE=75018719; PubMed=4138095;
RA Ruoslahti E., Pitko H., Vaheri A., Seppala M., Virolainen M.,
RA Kontinen A.;
RT "Alpha fetoprotein: structure and expression in man and inbred mouse
strains under normal conditions and liver injury.";
RL Johns Hopkins Med. J. Suppl. 3:249-255(1974).
RN [10]
RP GENE STRUCTURE.
RX MEDLINE=85182629; PubMed=2580830;
RA Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G.,
RA Tanacki T.;
RT "The human alpha-fetoprotein gene. Sequence organization and the 5'
flanking region.";
RL J. Biol. Chem. 260:5055-5060(1985).
RN [11]
RP METAL-BINDING.
RX MEDLINE=79001617; PubMed=80265;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Copper(II)-binding ability of human alpha-fetoprotein.";
RL Cancer Res. 38:3483-3486(1978).
RN [12]
RP BILIRUBIN-BINDING.
RX MEDLINE=80001710; PubMed=89900;
RA Aoyagi Y., Ikenaka T., Ichida F.;
RT "Alpha-Fetoprotein as a carrier protein in plasma and its bilirubin-
binding ability.";
RL Cancer Res. 39:3571-3574(1979).
RN [13]
RP SULFATION.
RX MEDLINE=86042625; PubMed=2414772;
RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;
RT "Tyrosine sulfation of proteins from the human hepatoma cell line
HepG2.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).
RN [14]
RP FUNCTION: Binds copper, nickel, and fatty acids as well as, and
bilirubin less well than, serum albumin. Only a small percentage
(less than 2%) of the human AFP shows estrogen-binding properties.
CC - SUBUNIT: Dimeric and trimeric forms have been found in addition
to the monomeric form.
CC - SUBCELLULAR LOCATION: Secreted.
CC - TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and
yolk sac.
CC - DEVELOPMENTAL STAGE: Occurs in the plasma of fetuses more than 4
weeks old, reaches the highest levels during the 12th-16th week of
gestation, and drops to trace amounts after birth. The serum level
in adults is usually less than 40 ng/ml. Afp occurs also at high
levels in the plasma and ascitic fluid of adults with hepatoma.
CC - PTM: Independent studies suggest heterogeneity of the amino-
terminal sequence of the mature protein and of the cleavage site
of the signal sequence.
CC - PTM: Sulfated.
CC - SIMILARITY: Belongs to the ALB/AFP/VDB family.
CC - SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; M10949; AA51674.1; -;
EMBL; M10950; AA51675.1; -;
EMBL; V01514; CAA24758.1; -;
EMBL; M16110; AAB58754.1; -;
EMBL; BC027881; AAH27881.1; -;
EMBL; Z19532; CAA79592.1; -;
PIR; A26624; FPHU.
HGSP; P02768; I578.
GlycoSiteDB; P02771; -;
Sienna-2DPAGE; P02771; -;
GeneW; HGNC:317; AFP.
MIN; 104150; -;
InterPro; IPR000264; Serum_albumin.
Pfam; PF00273; transport_prot; 3.
PRINTS; PR00802; SERUMALBUMIN.
ProDom; PD002486; Serum_albumin; 1.
SMART; SMO0103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 2.
Kw Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
Signal; Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 609 ALPHA-FETOPROTEIN.
FT DOMAIN 20 205 ALBUMIN 1.
FT DOMAIN 212 397 ALBUMIN 2.
FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 22 22 COPPER AND NICKEL.
FT DISULFID 99 114
FT DISULFID 113 124
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303

TISSUE= liver;
Cassidy A.I., Salklild C.K., Baverstock P., Wallace J.C.;
Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE OF 1-28 FROM N.A.
MEDLINE=83161037; PubMed=6187737;
Hache R.J.G., Wikocil R., Vasa M., Roy R.N., Lau P.C.K., Dealey R.G.;
"The 5' noncoding and flanking regions of the avian very low density
apolipoprotein II and serum albumin genes. Homologies with the egg
white protein genes." J. Biol. Chem. 258:4556-4564 (1983).
[3]
SEQUENCE OF 19-30.
MEDLINE=78019943; PubMed=911327;
Rosen A.M., Geller D.M.;
"Chicken microsomal albumin: amino terminal sequence of chicken
proalbumin." Biochem. Biophys. Res. Commun. 78:1060-1066 (1977).
[4]
ALLEGHENIC PROPERTIES.
MEDLINE=21381307; PubMed=11488669;
Quirce S., Maranon P., Umpierrez A., de las Heras M.,
Fernandez-Caldas E., Sastre J.;
"Chicken serum albumin (Gal d 5*) is a partially heat-labile inhalant
and food allergen implicated in the bird-egg syndrome." Allergy 55:754-762 (2001).
-!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
-!- SUBCELLULAR LOCATION: Secreted.
-!- TISSUE SPECIFICITY: Plasma.
-!- ALLEGEN: Causes an allergic reaction in human. Binds Ige.
Partially heat-labile allergen that may cause both respiratory and
food-allergy symptoms in patients with the bird-egg syndrome.
-!- SIMILARITY: Belongs to the ALB/APF/VDB family.
-!- SIMILARITY: Contains 3 albumin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; X60688; CAA43098.1; -;
EMBL; V00381; CAA23680.1; -;
PIR; S15571; ABCBS.
HSSP; P02758; 1E7B.
InterPro; IPR000264; Serum_albumin.
Pfam; PF00273; transprot_prot; 3.
PRINTS; P00802; SERUMALBUMIN.
ProDom; PD002486; Serum_albumin; 1.
SMART; SM00103; ALBUMIN; 3.
PROSITE; PS00212; ALBUMIN; 3.
Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
SIGNAL 1 18
PROPEP 19 23 SERUM ALBUMIN.
FT CHAIN 24 615
FT DOMAIN 24 209 ALBUMIN 1.
FT DOMAIN 216 401 ALBUMIN 2.
FT DOMAIN 408 599 ALBUMIN 3.
FT METAL 30 30 COPPER (BY SIMILARITY).
FT METAL 80 89 BY SIMILARITY.
FT DISULFID 102 118 BY SIMILARITY.
FT DISULFID 117 128 BY SIMILARITY.
FT DISULFID 152 197 BY SIMILARITY.
FT DISULFID 196 205 BY SIMILARITY.
FT DISULFID 228 274 BY SIMILARITY.
FT DISULFID 273 281 BY SIMILARITY.
FT DISULFID 293 307 BY SIMILARITY.
FT DISULFID 306 317 BY SIMILARITY.

[4] SEQUENCE OF 99-516 FROM N.A.
RP MEDLINE=88216123; PubMed=2452956;
RA Minghetti P.P., Law S.W., Dugaiczyk A.;
RT "The rate of molecular evolution of alpha-fetoprotein approaches that
of pseudogenes";
RL Mol. Biol. Evol. 2:347-358(1985).
RN [5]
RP SEQUENCE OF 477-551 FROM N.A.
RC STRAIN=BALB/c;
RA MEDLINE=90269606; PubMed=1971802;
RX Boccaccio C., Dechatrette J., Meunier-Rotival M.;
RT "Empty and occupied insertion site of the truncated LINE-1 repeat
located in the mouse serum albumin-encoding gene";
RL Gene 88:181-186(1990).
RN [6]
RP SEQUENCE OF 25-44.
RC TISSUE=Liver;
RX MEDLINE=9312044; PubMed=1286668;
RA Giometti C.S., Taylor J., Tollaksen S.L.;
RT "Mouse liver protein database: a catalog of proteins detected by two-
dimensional gel electrophoresis";
RL Electrophoresis 13:970-991(1992).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ011413; CAA09617.1; -;
DR EMBL; AK010025; BAB26650.1; -;
DR EMBL; BC049571; AAH49971.1; -;
DR EMBL; M16111; AAA37190.1; -;
DR EMBL; X13060; CAA31458.1; -;
DR PIR; A05139; A05139.
DR HSSP; P02768; 1E7B.
DR SWISS-2DPAGE; P07724; MOUSE.
DR MGD; MGI:87991; Alb1.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608
FT DOMAIN 25 205
FT DOMAIN 212 397
FT DOMAIN 404 595
FT METAL 27 27
FT METAL 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 269 277
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
BY SIMILARITY.

FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT CONFLICT 27 27 H -> D (IN REF. 6).
FT CONFLICT 33 33 H -> D (IN REF. 6).
FT CONFLICT 41 41 Q -> I (IN REF. 6).
SQ SEQUENCE 608 AA; 68692 MW; 2927F7C7EED3A61B4 CRC64;

Query Match 76.6%; Score 2378; DB 1; Length 608;
Best Local Similarity 72.4%; Pred. No. 2.6e-145;
Matches 422; Conservative 80; Mismatches 81; Indels 0; Gaps 0;

QY 1 DAKSEVAHRFKDLGEENFKALVLIATAQYLOOCPEPDHVKLVNEVTEFAKTCVADESAAE 60
DB 25 EAHKSEIAHRYNDLGEQHFGLVLIATFQYLOKSYDEHAKLVQEVTFDAKTCVADESAA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVLRPEV 120
DB 85 NCDKSLHTLFGDKLCAIPNLRENYGELADCCCTQEPERNECFLOHKDDNPSPFPERPEA 144
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQADKAACLLP 180
DB 145 EAMCTSPKENPTTFMGHYLHEVARRRHPYFYAPPELLYAEQYNEILTQCCAEADKESCLTP 204
QY 181 KLDELDRDEGKASAKQRLKASLOKGERAFKAWAVARLSQRPKAEFAEAVSKLVDTLTK 240
DB 205 KLDGVKEKALVSSVRQMKCSSMQKTFGERAFKAWAVARLSQTFPNADFAITKLATDLTK 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCECKPILLEKSHCIAEVENDEMPA 300
DB 265 VNKECCHGDLLECCADRAELAKYCNENQAISLKQTCCKPILKKAHCLSEVEDTMPA 324
QY 301 DLPSEADAFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLILAKTYETTLKEC 360
DB 325 DLPAIAADFVEDQVCKNYAEAKDVFLGTFLYYSRHPDYSVSLLLRLAKKYEATLEK 384
QY 361 CAADAPHECAKVPDEFKPLVERPQNLKONCELFEOLEKFKQNALLVRYTKKVPQVST 420
DB 385 CREANPPACVGTVLAEPQPLVEEPKLVNCTDLYEKLGEVQNALVRYTKAPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKHPEAKRMPCABDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
DB 445 PTLVEAARNLGRVGTCKCTLPEDQRLPCVEDYLSAILNRVCLLHEKTPVSEHVTCKCSGS 504
QY 481 LVNRPFCFSALEVDVETVVPKEFNAETTFHADICTLSEKERQIKKQATALVELVKHKEPAT 540
DB 505 LVERRPCFSAVTVDVETVVPKEFKAETTFHSDICTLPEKEKQIKKQATALVELVKHKEPAT 564
QY 541 KEQLKAVMDDPAFAFVEXCKKADDKETCFABEGKLVAAASQAAL 583
DB 565 AEQLKTVMDPFAFLDTCCKAADKDTCTFSTEGPNLVTRCKDAL 607

RESULT 13
ID ALBU CHICK STANDARD; PR7; 615 AA.
AC P19121;
DT 01-NOV-1990 (Rel. 15, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor (Alpha-livetin) (Allergen Gal d 5).
ALB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RN NCBI_TaxID=9031;
RP SEQUENCE FROM N.A.

DR	PIR; JC5838; JC5838.	DB	507	VERRPCFSALPVDETVVPKFEKAEFTTHANICTLPEKEKQMEKOTALAEVLKHKPQATE	566
DR	HSP; P02768; 1E7B.	QY	542	EOLKAVMDFFAFAFVEKCKKADKXETCFABEKGKLVAAASQAL	583
DR	InterPro; IPR000264; Serum albumin.	DB	567	EOLKVMGDFABFLEKCKQEDKACFSTEGPKLVAESQKAL	608
DR	Pfam; PF00273; transport pT0t; 3.				
DR	PRINTS; PR00802; SERUMALBUMIN.				
DR	ProDom; PD002486; Serum albumin; 1.				
DR	SMART; SM00103; ALBUMIN; 3.				
DR	PROSITE; PS00212; ALBUMIN; 3.				
KW	Metal-binding; Lipid-binding; Repeat; Signal; Copper.				
FT	SIGNAL 1 18	ALBU_MOUSE			
FT	PROPEP 19 24	AC	P07724; Q61802;	STANDARD; PRT; 608 AA.	
FT	CHAIN 25 609	DT	01-APR-1988 (Rel. 07, Created)		
FT	DOMAIN 25 206	DT	15-JUL-1999 (Rel. 38, Last sequence update)		
FT	DOMAIN 213 398	DT	15-MAR-2004 (Rel. 43, Last annotation update)		
FT	DOMAIN 405 596	DE	Serum albumin precursor.		
FT	METAL 28 87	GN	ALBU OR ALB1 OR ALB-1.		
FT	DISULFID 78 87	OS	Mus musculus (Mouse).		
FT	DISULFID 100 116	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
FT	DISULFID 115 126	OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
FT	DISULFID 149 184	OX	NCBI_TaxID=10090;		
FT	DISULFID 193 202	RN	[1]		
FT	DISULFID 225 271	RP	SEQUENCE FROM N.A.		
FT	DISULFID 270 278	RP	TISSUE=Liver;		
FT	DISULFID 303 314	RC	van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;		
FT	DISULFID 341 366	RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.		
FT	DISULFID 365 394	RP	SEQUENCE FROM N.A.		
FT	DISULFID 417 463	RC	STRAIN=CS7BL/6J; TISSUE=Tongue;		
FT	DISULFID 462 473	RC	MEDLINE=21085660; PubMed=11217851;		
FT	DISULFID 486 502	RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
FT	DISULFID 501 512	RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
FT	DISULFID 539 584	RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
FT	DISULFID 583 592	RA	Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,		
FT	SEQUENCE 609 AA; 68940 MW; 9CA5F97F67EFLA48 CRC64;	RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant I.,		
		RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
		RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
		RA	Schirml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,		
		RA	Sakai K., Okido T., Furuno M., Rono H., Baldarelli R., Barsh G.,		
		RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
		RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
		RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
		RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
		RA	Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
		RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
		RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
		RA	Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,		
		RA	Hayashizaki Y.;		
		RT	"Functional annotation of a full-length mouse cDNA collection.;"		
		RL	Nature 409:685-690(2001).		
		RP	[3]		
		RP	SEQUENCE FROM N.A.		
		RC	STRAIN=FVB/N; TISSUE=Liver;		
		RC	MEDLINE=22388257; PubMed=12477932;		
		RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
		RA	Klausner R.D., Collins F.S., Wagner L., Schaefer C.F., Schuler G.D.,		
		RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
		RA	Diatchenko L., Jordan H., Moore T., Max S.I., Wang J., Hoehle F.,		
		RA	Hopkins R.F., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
		RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
		RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,		
		RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
		RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
		RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
		RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
		RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		
		RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		
		RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
		RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
		RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,		
		RA	Schmurch A., Schein J.E., Jones S.J.M., Marra M.A.;		
		RT	"Generation and initial analysis of more than 15,000 full-length		
		RT	human and mouse cDNA sequences.;"		
		RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		

QY	2	AHSEVAHPRKDIIGENFKALVLIAPAOYLQCPPEHDKLVNEVTEFAKTCVADESAN	61
DB	27	AHSELAHRYKDIIGKFKGLVLYTSQVLOKCSYEHVKLVREVTDFASNAKDESAN	86
QY	62	CDKSLHTLFGDKLCTVATLETYEMADCCAKQEPERNESCFLQHKDDNNLRLVRPEVD	121
DB	87	CDKSLHTLFGDKLCSLPNFKYAEAMADCCAKQEPERNESCFLQHKDDNNLRLVRPEVD	146
QY	122	VMCTAFHDNETFLKLYLVEIARRHPVYVAPPELLFAKRYKAAFTCCCOADKAACLLPK	181
DB	147	AMCTAFQENAEAFNGHYLHEVARHPYFGPELLYLADKTAVLTCCADKAGCLLPK	206
QY	182	LDELRDGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPFAEFVSKLVTDLTKV	241
DB	207	LDALKERALYSAYVQRLKCSMMKFGERAFKAWAVARMSQTFPNADFAETKLTATLTKV	266
QY	242	HTECHGDLLECADRADLAKYICENQDSISSKLKCEKPELEKSHCAEVENDEMPAD	301
DB	267	TQECCHGDLLECADRADLAKYICENQDSISSKLKCEKPELEKSHCAEVENDEMPAD	326
QY	302	LPSLAADFVSKDYCKNYAKADKDFGLGWFLEYEARHPDYSVVLLELAKTYETTLKCC	361
DB	327	LPALTAFVDEKDYCKNYAKADKDFGLGWFLEYEARHPDYSVVLLELAKTYETTLKCC	386
QY	362	AAADPHSCYAKVFDEKPLVEEPQNLKONCELPQGEYKFNALLVRYTKKQVQSTP	421
DB	387	AEADPHACVGHVDFEKLVEEPQNLKONCELPQGEYKFNALLVRYTKKQVQSTP	446
QY	422	TLVSVNRNLGVKCKCKGHPKAMPKCAEDYLSVVLNQLCVLHEKTEPVSQVTKCTTESL	481
DB	447	TLVSVNRNLGVKCKCKGHPKAMPKCAEDYLSVVLNQLCVLHEKTEPVSQVTKCTTESL	506
QY	482	VNRPCFSALVDETVVPKFEKAEFTTHANICTLPEKEKQMEKOTALAEVLKHKPQATE	541

RC TISSUE=Liver;
 RX MEDLINE=89016582; PubMed=3174440;
 RA Baldwin G.S., Weinstock J.;
 RT "Nucleotide sequence of porcine liver albumin.";
 RL Nucleic Acids Res. 16:9045-9045(1988).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X12422; CAA30970.1; -;
 CC EMBL: M36787; AAA30988.1; -;
 CC PIR: S01382; ABPGS.
 CC HSP: P02768; IEVH.
 CC InterPro: IPR000264; Serum albumin.
 CC Pfam: PF00273; transport prot; 3.
 CC PRINTS: PR00802; SERUMALBUMIN.
 CC ProDom: PD002486; Serum albumin; 1.
 CC SMART: SM00103; ALBUMIN_3.
 CC PROSITE: PS00212; ALBUMIN; 3.
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 KW NON_TER 1
 FT SIGNAL <1 16 BY SIMILARITY.
 FT PROPEP 17 22 BY SIMILARITY.
 FT CHAIN 23 605 SERUM ALBUMIN.
 FT DOMAIN 23 202 ALBUMIN 1.
 FT DOMAIN 209 394 ALBUMIN 2.
 FT DOMAIN 401 592 ALBUMIN 3.
 FT METAL 31 31 COPPER (BY SIMILARITY).
 FT DISULFID 75 84 BY SIMILARITY.
 FT DISULFID 97 113 BY SIMILARITY.
 FT DISULFID 112 123 BY SIMILARITY.
 FT DISULFID 145 190 BY SIMILARITY.
 FT DISULFID 183 198 BY SIMILARITY.
 FT DISULFID 221 267 BY SIMILARITY.
 FT DISULFID 266 274 BY SIMILARITY.
 FT DISULFID 286 300 BY SIMILARITY.
 FT DISULFID 299 310 BY SIMILARITY.
 FT DISULFID 337 382 BY SIMILARITY.
 FT DISULFID 381 390 BY SIMILARITY.
 FT DISULFID 413 459 BY SIMILARITY.
 FT DISULFID 458 469 BY SIMILARITY.
 FT DISULFID 482 498 BY SIMILARITY.
 FT DISULFID 497 508 BY SIMILARITY.
 FT DISULFID 535 580 BY SIMILARITY.
 FT DISULFID 579 588 BY SIMILARITY.
 FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).
 FT SEQUENCE 605 AA; 69410 MW; 3E556B0DD1A1FAFF CRC64;
 Query Match 77.7%; Score 2411.5; DB 1; Length 605;
 Best Local Similarity 76.0%; Pred. No. 1.9e-147;
 Matches 438; Conservative 67; Mismatches 70; Indels 1; Gaps 1;
 QY 1 DAKSEVAHFKDGLGEENFVLAFAVQLQCPEDHVKLVNEVTEFAKTCVADSAE 60
 DB 23 DTYKSEIAHFKDGLGEYFGLVLAFAVSHQLQCPVEEHVKLVREVTEFAKTCVADSAE 82
 QY 61 NCDKSLTFLFGDKLCTVATIREYGVEMADCCAKQEPERNECFLOHKKDDPNLPRLVREPV 120
 DB 83 NCDKSLTFLFGDKLCTVATIREYGVEMADCCAKQEPERNECFLOHKKDDPNLPRLVREPV 141

QY 121 DVNCTAFHDNBEFTLKKVLYELIARRHPYFAPPELLFFAKRYKAAFTCCQADKAACLLP 180
 DB 142 VALCADFOEQKFWGKYLYEIAKHPYFAPPELLYYAIYKDVFESECCQADKAACLLP 201
 QY 181 KDELDRDEKASSAKORLKCASLQKQGERAFKAMAVARLSORFFPKAEFAEYVKLYDTLTK 240
 DB 202 KIEHLREKVLTSAAKORLKCASIQKQGERAFKAWLSORFFPKADFTISKIVTDLAK 261
 QY 241 VITECHGDLLECADRADLAKYICENQDSISSKULKECEKPLLEKSHCIAEVENDEMPA 300
 DB 262 VHKECHGDLLECADRADLAKYICENQDTISTKLKCECDKLEKSHCIAEAKDELP 321
 QY 301 DIPSLAADFVSKDKVCKYAEAKDFLGMFLVEYARRHPDYVWLLLRLLAKIYETTLK 360
 DB 322 DNLPHSHDFVEKVKCKYKEAKDVLGTLFVYSRRHPDYVSLLLRIAKIYEATLEDC 381
 QY 361 CAAADPHECYAKVDFEFPKPLVEEPQNLKQNCLEFEQLEGEYFONALLVRYTKKPVQVST 420
 DB 382 CAKEDPPACYATVDFKQPLVDKPNLKQNCLEFEKLEGEYFONALLVRYTKKPVQVST 441
 QY 421 PTLVEYSRNLGKVGSKCKHPSAKMPCAEEDVLSVLNOLCVLHEKTPVSDRVTKCCTES 480
 DB 442 PTLVEVARKLGLVSKCKRPEERUSCAEDVLSVLNOLCVLHEKTPVSEKVKCCTES 501
 QY 481 LVNRPCFSALVDETYTPKFNATFTTFHADICTLSEKEROIKKQATALVELVGHKPKAT 540
 DB 502 LVNRPCFSALTPDETYKPEFVEGTFTFHADICTLPEDEKQIKKQATALVELVGHKPKAT 561
 QY 541 KEQLKAWMDFFAIVEKCKKADKCTCFABEGKLV 576
 DB 562 EEQLTTLVGNFAAFVQKCAAPDHEACFAVEGPKFV 597
 RESULT 11
 ALBU_MERUN STANDARD; PRT; 609 AA.
 AC O35090;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB
 OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
 OC Meriones.
 OC NCBI_TaxID=10047;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGS IDR; TISSUE=Liver;
 RX MEDLINE=98116663; PubMed=9455485;
 RA Yoshida K., Seto-Oshihama A., Sinohara H.;
 RT "Sequencing of cDNA encoding serum albumin and its extrahepatic
 RL synthesis in the Mongolian gerbil, Meriones unguiculatus.";
 RL DNA Res. 4:351-354(1997).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AB006197; BAA21765.1; -;
 DR

RN SEQUENCE OF 25-222.
 RX MEDLINE=78109429; PubMed=564345;
 RA Isemura S., Ikenaka T.;
 RT "Amino acid sequences of fragments I and II obtained by cyanogen
 bromide cleavage of rat serum albumin.";
 RL J. Biochem. 83:35-48(1978).
 RN [4]
 RP SEQUENCE OF 223-288 AND 572-608.
 RX MEDLINE=76260153; PubMed=956149;
 RA Isemura S., Ikenaka T.;
 RT "Fragmentation of rat serum albumin by cyanogen bromide cleavage and
 the amino acid sequences of four fragments.";
 RL J. Biochem. 79:1183-1196(1976).
 RN [5]
 RP SEQUENCE OF 166-174.
 RX TISSUE=Plasma;
 RA MEDLINE=87194805; PubMed=2437111;
 RA Caraway R.E., Mitra S.P., Cochran D.E.;
 RT "Structure of a biologically active neurotensin-related peptide
 obtained from pepsin-treated albumin(s).";
 RL J. Biol. Chem. 262:5968-5973(1987).
 RN [6]
 RP COPPER-BINDING.
 RX MEDLINE=79001617; PubMed=80265;
 RA Aoyagi Y., Ikenaka T., Ichida F.;
 RT "Copper(II)-binding ability of human alpha-fetoprotein.";
 RL Cancer Res. 38:3483-3486(1978).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -1- FUNCTION: NRP regulates fat digestion, lipid absorption, and
 blood flow (Potential).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.
 CC -1- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V01222; CAA24532.1; -.
 CC PIR; A93872; ABRTS.
 CC HSP; P02768; 1E7B.
 CC InterPro; IPR00264; Serum albumin.
 CC Pfam; PF00273; transport_prot; 3.
 CC PRINTS; PR00802; SERUMALBUMIN.
 CC ProDom; PD002486; Serum_albumin; 1.
 CC SMART; SM00103; ALBUMIN; 3.
 CC PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT PEPTIDE 166 174
 FT PEPTIDE 166 174
 FT DOMAIN 25 205
 FT DOMAIN 212 397
 FT DOMAIN 404 595
 FT METAL 27 27
 FT METAL 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 148 193
 FT DISULFID 192 201
 FT DISULFID 224 270
 FT DISULFID 269 277
 FT DISULFID 289 303

FT DISULFID 302 313 BY SIMILARITY.
 FT DISULFID 340 385 BY SIMILARITY.
 FT DISULFID 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.
 FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT VARIANT 262 262
 FT CONFLICT 174 174 Y -> L (IN REF. 5).
 SQ SEQUENCE 608 AA; 68718 MW; 5BB497A282411AB7 CRC64;
 Query Match 78.2%; Score 2426; DB 1; Length 608;
 Best Local Similarity 73.4%; Pred. No. 2.2e-148;
 Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;
 QY 1 DAHSEVAHPRKDLGHNFKALVLIAPQVLOCPEDHVKLVNVEYTEPAKTCAVDESAE 60
 DB 25 EAHKSEIAHPRKDLGHNFKALVLIAPQVLOCPEDHVKLVNVEYTEPAKTCAVDESAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPEV 144
 QY 121 DVMTAFHNEETPLKYLVEIARRHPYFYAPELLFFAKRYKAAFTCCCAADKAAALLP 180
 DB 145 EAMCTSPQENPTSLGHVLEHVARHPYFYAPELLFFAKRYKAAFTCCCAADKAAALLP 204
 QY 181 KLDELREGKASSAKQRLKQKGERAFKAWAVARLSQRFPKAEFAEVSKLVDLTG 240
 DB 205 KLDVKEKALVAARQRMKSSQWQGERAFKAWAVARLSQRFPKAEFAEVSKLVDLTG 264
 QY 241 VHTCCGHDLLCADDRAADLAKYICENQDSISKLKCECKPILKSHCHIAEVENDEMPA 300
 DB 265 INKECHGDLLECCADRAELAKYICENQDSISKLKCECKPILKSHCHIAEVENDEMPA 324
 QY 301 DLPSLAADFVESKVCNVAEAKDVFLGMPLYEVARHPDYSVVLRLAKTVEITTEK 360
 DB 325 DLPSLAADFVESKVCNVAEAKDVFLGMPLYEVARHPDYSVVLRLAKTVEITTEK 384
 QY 361 CAADAPHECAKVPDEPKLVPEEQNLKONCELFQGLGKQFQNALLVRYTKKPVQVST 420
 DB 385 CAEGDPACVTVLAEPQLVPEEQNLKONCELFQGLGKQFQNALLVRYTKKPVQVST 444
 QY 421 PTLVEVSRNLGKVGCKCKHPEAKRMPCAEYLSVNLQVLEKTPVSDRVTKCCTES 480
 DB 445 PTLVEVSRNLGKVGCKCKHPEAKRMPCAEYLSVNLQVLEKTPVSDRVTKCCTES 504
 QY 481 LVNRRPCFSALVEDETVYKPEFNAETFTFHADICTLSEKERQIKKQATALVELVHKPKAT 540
 DB 505 LVNRRPCFSALVEDETVYKPEFNAETFTFHADICTLSEKERQIKKQATALVELVHKPKAT 564
 QY 541 KEQLKAVMDDFAAFVEKCKKADDETCFAEKGKLVAAASQAL 583
 DB 565 EDQLKTVMGDFQVDFVCKCKAADKDNCFATEGNLVARSKAL 607
 RESULT 10
 ID ALBU_PIG STANDARD; PRT; 605 AA.
 AC P08835; Q29018;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor (Fragment).
 GN ALB.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.

RESULT 8
ALBU SHEEP STANDARD; PRT; 607 AA.
AC P14639.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=9009888; PubMed=2602160;
RX Brown W.M., Diegisielska K.M., Foreman R.C., Saunders N.R.;
RT "Nucleotide and deduced amino acid sequence of sheep serum albumin.";
RL Nucleic Acids Res. 17:10495-10495(1989).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17055; CAA34903.1; -.
DR PIR; S06936; ABSH.
DR HSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; P00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24 BY SIMILARITY.
FT CHAIN 25 607 SERUM ALBUMIN.
FT DOMAIN 25 204 ALBUMIN 1.
FT DOMAIN 211 396 ALBUMIN 2.
FT DOMAIN 403 594 ALBUMIN 3.
FT METAL 27 27 COPPER (BY SIMILARITY).
FT DISULFID 77 86 BY SIMILARITY.
FT DISULFID 99 115 BY SIMILARITY.
FT DISULFID 114 125 BY SIMILARITY.
FT DISULFID 147 192 BY SIMILARITY.
FT DISULFID 191 200 BY SIMILARITY.
FT DISULFID 223 269 BY SIMILARITY.
FT DISULFID 268 276 BY SIMILARITY.
FT DISULFID 288 302 BY SIMILARITY.
FT DISULFID 301 312 BY SIMILARITY.
FT DISULFID 339 384 BY SIMILARITY.
FT DISULFID 383 392 BY SIMILARITY.
FT DISULFID 415 461 BY SIMILARITY.
FT DISULFID 460 471 BY SIMILARITY.
FT DISULFID 484 500 BY SIMILARITY.
FT DISULFID 499 510 BY SIMILARITY.
FT DISULFID 537 582 BY SIMILARITY.
FT DISULFID 581 590 BY SIMILARITY.
SQ SEQUENCE 607 AA; 69188 MW; 84979A87F8B86596 CRC64;

Query Match 78.4%; Score 2432.5; DB 1; Length 607;
Best Local Similarity 75.0%; Pred. No. 8.4e-149;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;
QY 1 DAHSEVHRPKDGLGEENFKALVLTAPAYLOQCPEPDHVKLVNVEVTEFAKTCVADESAAE 60
DB 25 DTHKSEAHRFNDLGEENFQGLVLTAFSQYLOQCPEPDHVKLVNVEVTEFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQKPERNECFLOHKDDNPRLVPRVSV 120
DB 85 GCDKSLHTLFGDELCKVATLRETYGDMADCCQKPERNECFLNHKDDSPDLPKL-KPEP 143
QY 121 DVNCTAFHNEETFLKKYLVEIARRHPYAPPELLFPAKRYKAAAFTECCQADKACCLIP 180
DB 144 DTLCAEFKADKKFWGKYLVEYARRHPYAPPELLYANKYNGVFOCCQADKACCLIP 203
QY 181 KLDELDEGKASSAKORLKCASLOKFGERAFAKAWAVARLSQRPFPKAEFVSKLVTDLTK 240
DB 204 KIDAMREKVLASSARQRLCASIQKFGERALKAWVARLSQKFPKADFTDVKIVTDLTK 263
QY 241 VHTCCGDLLECCADDRADLAKYICENQDSISGKKECCKEKPLLEKSHGICIAEVNDMPA 300
DB 264 VHKCCGDLLECCADDRADLAKYICDHQDALSSKKECCDKPVLEKSHGICIAEVDKAVPE 323
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFVEYARRHPDYSVLLLRLLAKTVETLEKC 360
DB 324 NLPLTADFAEDKEVKNYQEAKDVLGSLFVEYSRRHPYAVSVLLRLAKEVEATLEDC 383
QY 361 CAAADPHECYAKVDFBPKPLVEEPQNLIKONCELPQQLGEYKFNQALLVRYTKKVPQVST 420
DB 384 CAKEDPHACVATVDFDKLHLVDPEQNLIKONCELPQQLGEYKFNQALLVRYTKKVPQVST 443
QY 421 PTLVEYSRLGKVGSKCKHPEAKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480
DB 444 PTLVEYSRLGKVGSKCKHPEAKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 503
QY 481 LVNRRPCFSALYDETVYVPKFNFAETTFHADICTLSEKERQIKKOTALVELVKKPKAT 540
DB 504 LVNRRPCFSLTLDETVYVPKFPDEKFTTFHADICTLPTDEKQIKKOTALVELVKKPKAT 563
QY 541 KEQLKAVMDFFAFAVEKCCCKADDEKTCFAEGKLVAAASQAL 583
DB 564 DEQLKVTMENFVAFVDKCAADDEKGCFLVLEGPRLVASTQAL 606
RESULT 9
ALBU RAT
ID -ALBU RAT STANDARD; PRT; 608 AA.
AC P02770; P11382;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor [Contains: Neurotensin-related peptide (NRP)].
GN ALB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81223722; PubMed=7017712;
RA Sargent T.D., Yang M., Bonner J.;
RT "Nucleotide sequence of cloned rat serum albumin messenger RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:243-246(1981).
RN [2]
RP SEQUENCE OF 1-38, AND PROCESSING.
RX MEDLINE=77249657; PubMed=893447;
RA Strauss A.W., Bennett C.D., Donohue A.M., Rodkey J.A., Alberts A.W.;
RT "Rat liver pre-proalbumin: complete amino acid sequence of the pre-
RT piece. Analysis of the direct translation product of albumin
RT messenger RNA.";
RL J. Biol. Chem. 252:6846-6855(1977).

Db 144 NTLDEPKADKKFWGKLYEIAIRRHYPFYAPPELLYVANKYNGVFQCCQAEKDGACLLP 203
QY 181 KLDELDEGKASSAKQRLKASLOKQGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240
Db 204 KIETWREKVLASSARQRLRCASIQKQGERALKAWAVARLSQKPKAEFVETKLVTLTK 263
QY 241 VHTCCCHGDLLECADDRLADLAKYICENODSISSKKECCPKLEKSHCAEVENDEMPA 300
Db 264 VKHCCHGDLLECADDRLADLAKYICDNQDTISSKKECCPKLEKSHCAEVEKOAPE 323
QY 301 DLPSLAADFVSKVKCKNYAEAKDVFGLGMFLYEVARRHPDYVSVLLRLAKTYETTLK 360
Db 324 NLPPLTADFAEDKVKCKNYQAKDAFLGSFLYEVARRHPDYVSVLLRLAKTYETTLK 383
QY 361 CAADDPHECYAKVDFEKPVEEQNLKQNCCELFEQGLGEYKQFONALLVRYTKVPOYST 420
Db 384 CAKDDPHACYSTVDFKHLVDEPNLQKNCDOFEKLGEGYFONALLVRYTKVPOYST 443
QY 421 PTLVEVSRLGKVGSKCCPKPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 444 PTLVEVSRLGKVGSKCCPKPEAKRMPCTEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 503
QY 481 LVNRRPCFSALEVDYTPK3FNAETTFHADICTLSEKROIKQKOTALVELVHKPKAT 540
Db 504 LVNRRPCFSALEVDYTPK3FNAETTFHADICTLSEKROIKQKOTALVELVHKPKAT 563
QY 541 KEQLKAVMDDFAAFVEKCKKADDEKTCFAEKGKLVAAQAAL 583
Db 564 BEQLKAVMDDFAAFVEKCKKADDEKTCFAEKGKLVAAQAAL 606

RESULT 7
ID ALBU RABIT STANDARD; PRT; 608 AA.
AC P49065;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; U18344; AAB58347.1; -.
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum albumin.
DR Pfam; PF00273; transport prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00012; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL 1 18
FT PROPEP 19 24
FT CHAIN 25 608
FT DOMAIN 25 205
FT DOMAIN 212 397
FT DOMAIN 404 595
FT METAL 27 27
FT DISULFID 77 86
FT DISULFID 99 115
FT DISULFID 114 125
FT DISULFID 148 193
FT DISULFID 192 201
FT DISULFID 224 270
FT DISULFID 289 303
FT DISULFID 302 313
FT DISULFID 340 385
FT DISULFID 416 462
FT DISULFID 461 472
FT DISULFID 485 501
FT DISULFID 500 511
FT DISULFID 538 583
FT DISULFID 582 591
SQ SEQUENCE 608 AA; CF5E92647AAPE9A2 CRC64;
Query Match 78.8%; Score 2446; DB 1; Length 608;
Best Local Similarity 74.1%; Pred. No. 1.1e-149; Indels 0; Gaps 0;
Matches 433; Conservative 77; Mismatches 74;
QY 1 DAHKEVAHRFKDGLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 EAHKEIAHRFNDVGEHFIQGLVITFSQYKQCPYEEHAKLVKEVTDLAKAADAESA 84
QY 61 NCDKSLHTLFGDKLCTVATLAEITYGEMADCAKQEPERNECEFLQKNDPNLPLRVREV 120
Db 85 NCDKSLHTLFGDKLCTVATLAEITYGEMADCAKQEPERNECEFLQKNDPNLPLRVREV 144
QY 121 DVNCTAFHDNEETFLKLYEIAIRRHYPFYAPPELLYVANKYNGVFQCCQAEKDGACLLP 180
Db 145 DVLKRAFDHDEKAFGHVLYEVARRHPFYAPPELLYVANKYNGVFQCCQAEKDGACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKQGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240
Db 205 KLDELDEGKASSAKQRLKASLOKQGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 264
QY 241 VHTCCCHGDLLECADDRLADLAKYICENODSISSKKECCPKLEKSHCAEVENDEMPA 300
Db 265 VKHCCHGDLLECADDRLADLAKYICDNQDTISSKKECCPKLEKSHCAEVEKOAPE 324
QY 301 DLPSLAADFVSKVKCKNYAEAKDVFGLGMFLYEVARRHPDYVSVLLRLAKTYETTLK 360
Db 325 GLPAVAEEFVEDKVKCKNYEAKDLFLGKFLYEVARRHPDYVSVLLRLAKTYETTLK 384
QY 361 CAADDPHECYAKVDFEKPVEEQNLKQNCCELFEQGLGEYKQFONALLVRYTKVPOYST 420
Db 385 CATDDPHACYSTVDFKHLVDEPNLQKNCDOFEKLGEGYFONALLVRYTKVPOYST 444
QY 421 PTLVEVSRLGKVGSKCCPKPEAKRMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLGKVGSKCCPKPEAKRMPCTEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTPK3FNAETTFHADICTLSEKROIKQKOTALVELVHKPKAT 540
Db 505 LVNRRPCFSALEVDYTPK3FNAETTFHADICTLSEKROIKQKOTALVELVHKPKAT 564
QY 541 KEQLKAVMDDFAAFVEKCKKADDEKTCFAEKGKLVAAQAAL 584
Db 565 NDQLKTVGVEFTALLDKCCSAEDKACFAVEGPKLVSSKATLG 608

[2] SEQUENCE FROM N.A., AND VARIANT THR-214.
 RC TISSUE=Liver;
 RA Barry T., Power S., Gannon F.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 [3] SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RC Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
 RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 [4] SEQUENCE FROM N.A., AND VARIANT THR-214.
 RP Wu H.T., Huang M.C.;
 RA "The complete cDNA sequence of bovine serum albumin."
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 [5] SEQUENCE OF 1-32.
 RP MEDLINE=80024278; PubMed=489109;
 RA McGilivray R.T.A., Chung D.W., Davie E.W.;
 RL "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-terminal sequence of prealbumin."
 [6] Eur. J. Biochem. 98:477-485(1979).
 [7] SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
 RP Brown J.R.;
 RA "Structure of bovine serum albumin."
 RL Fed. Proc. 34:591-591(1975).
 [8] REVISIONS TO 190-195.
 RP Brown J.R.;
 RA Submitted (APR-1975) to the PIR data bank.
 [9] SEQUENCE OF 402-433.
 RP MEDLINE=82023364; PubMed=7283978;
 RA Reed R.G., Putnam F.W., Peters T. Jr.;
 RL "Sequence of residues 400-403 of bovine serum albumin."
 [10] Biochem. J. 191:867-868(1980).
 [11] SEQUENCE OF 19-28.
 RP MEDLINE=77134075; PubMed=843354;
 RA Patterson J.E., Geller D.M.;
 RL "Bovine microsomal albumin: amino terminal sequence of bovine prealbumin."
 [12] Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
 [13] SEQUENCE, AND REVISIONS TO 118-119 AND 180.
 RP MEDLINE=91083649; PubMed=2260975;
 RA Hirayama K., Akashi S., Furiya M., Fukuhara K.-I.;
 RL "Rapid confirmation and revision of the primary structure of bovine serum albumin by ESIMS and Frit-FAB LC/MS."
 [14] Biochem. Biophys. Res. Commun. 173:639-646(1990).
 [15] SEQUENCE OF 25-41.
 RP MEDLINE=88267456; PubMed=3389500;
 RA Heieh J.C., Lin F.P., Tam M.F.;
 RL "Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel: a preparative method for isolating proteins for N-terminal microsequencing."
 [16] Anal. Biochem. 170:1-8(1988).
 [17] SEQUENCE OF 437-451.
 RP Vilbois F.;
 RA Submitted (AUG-1998) to Swiss-Prot.
 [18] DISULFIDE BONDS.
 RP Brown J.R.;
 RA "Structure of serum albumin: disulfide bridges."
 RL Fed. Proc. 33:1389-1389(1974).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Plasma.
 CC -1- ALLERGEN: Causes an allergic reaction in human.
 CC -1- SIMILARITY: Belongs to the ALB/AFP/VDB family.
 CC -1- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL; M73993; AAA51411.1; -;
 DR EMBL; X58989; CAA41735.1; -;
 DR EMBL; Y17769; CAA76847.1; -;
 DR EMBL; AF542068; AAN17824.1; -;
 DR HSSP; P02768; 1E7B.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen;
 KW Polymorphism.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 607 SERUM ALBUMIN.
 FT DOMAIN 25 204 ALBUMIN 1.
 FT DOMAIN 211 396 ALBUMIN 2.
 FT DOMAIN 403 594 ALBUMIN 3.
 FT METAL 27 27 COPPER (BY SIMILARITY).
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 147 192
 FT DISULFID 191 200
 FT DISULFID 223 269
 FT DISULFID 268 276
 FT DISULFID 288 302
 FT DISULFID 301 312
 FT DISULFID 339 384
 FT DISULFID 383 392
 FT DISULFID 415 461
 FT DISULFID 460 471
 FT DISULFID 484 500
 FT DISULFID 499 510
 FT DISULFID 537 582
 FT DISULFID 581 590
 FT VARIANT 214 214
 FT CONFLICT 302 302 A -> T.
 FT CONFLICT 304 305 C -> K (IN REF. 6).
 FT CONFLICT 324 324 N -> D (IN REF. 6).
 FT CONFLICT 394 395 ST -> TS (IN REF. 6).
 FT CONFLICT 437 437 K -> R (IN REF. 12).
 FT CONFLICT 493 494 SE -> ES (IN REF. 6).
 SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;
 Query Match 79.0%; Score 2450.5; DB 1; Length 607;
 Best Local Similarity 75.8%; Pred. No. 5.9e-150;
 Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 1;
 QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAFQVLCQCPFDHVKLVNEVTEFAKTCVADSSAE 60
 Db 25 DTHKSEIHRFKDLGSEHFKGLVLIAPSVLIQCQPFDEHVKLVNELTEFAKTCVADSSHA 84
 QY 61 NCDKSLHTLFDKLCCTVATIRETYGEMADCCAKOEPERNECFLOHKDNNPLPLVPEV 120
 Db 85 GCEKSLHTLFDKLCVASTIRETYGEMADCCAKOEPERNECFLOHKDNNPLPLVPEV 143
 QY 121 DVMCTATHDNETFLKKVLYEIAARRHPYFVAPPELLFFAKRYKAAFTCCQAADKAACLLP 180

Db 505 LVNRPFCPSGLEVDYVYKFNATFTFHADLCTLPEAKQVKKQFALVELLKHKPKAT 564
 QY 541 KEQLKAVMDDPFAFVEKCKADKCTCFABEGKLVAAQAAL 583
 Db 565 DEQLKTVNGDFGAFVEKCAENKEGCFSEBGPKLVAQAAL 607

RESULT 5
 ALBU_HORSE
 ID ALBU_HORSE STANDARD; PRT; 607 AA.
 AC P35747;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serum albumin precursor (Allergen Equ c 3).
 GN ALB.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RC TISSUE=Liver;
 RX MEDLINE=93345495; PubMed=8344282;
 RA Ho J.X., Holowachuk E.W., Norton E.J., Twigg P.D., Carter D.C.;
 RT "X-ray and primary structure of horse serum albumin (Equus caballus)
 at 0.27-nm resolution.";
 RL Eur. J. Biochem. 215:205-212 (1993).
 CC -|- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water. Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- TISSUE SPECIFICITY: Plasma.
 CC -|- ALLERGEN: Causes an allergic reaction in human. Binds IGE.
 CC -|- SIMILARITY: Belongs to the ALB/AFP/VDB family.
 CC -|- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X74045; CAA52194.1; -
 CC PIR; S34053; ABHOS.
 CC HSP; P02768; 1E7B.
 CC InterPro; IPR000264; Serum albumin.
 CC Pfam; PF00273; transport prot; 3.
 CC PRINTS; PR00802; SERUMALBUMIN.
 CC PRODOM; PD00486; Serum_albumin; 1.
 CC SMART; SM00103; ALBUMIN; 3.
 CC PROSITE; PS00212; ALBUMIN; 3.
 CC Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
 KW SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 607
 FT DOMAIN 25 204
 FT DOMAIN 211 396
 FT DOMAIN 403 594
 FT METAL 27 27
 FT METAL 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 147 192
 FT DISULFID 191 200
 FT DISULFID 223 269
 FT DISULFID 268 276
 FT DISULFID 288 302
 FT DISULFID 301 312

FT DISULFID 339 384
 FT DISULFID 383 392
 FT DISULFID 415 461
 FT DISULFID 460 471
 FT DISULFID 484 500
 FT DISULFID 499 510
 FT DISULFID 537 582
 FT DISULFID 581 590
 SQ SEQUENCE 607 AA; 68598 MW; 256F6E830A1B90C5 CRC64;
 Query Match 79.8%; Score 2475.5; DB 1; Length 607;
 Best Local Similarity 76.3%; Pred. No. 1.5e-151;
 Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;
 QY 1 DAHSEVAHRFKDLGRENKALVLIARAFAYLOCPEDHVKLVNVEYTERAKTVADESAE 60
 Db 25 DTHKSEIAHRFNDLGEKFKGLVLFVAFSQYLOCPEDHVKLVNVEYTERAKTVADESAE 84
 QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 120
 Db 85 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPEV 143
 QY 121 DVMCTAFHNEETFLKYLVEIARRHPYFYAPPELLFFAKYKAAFTCCCOAADAACALLP 180
 Db 144 DAQCAAFQEDPDKFLGKLYEVARRRHPYFYGPPELLFFHAEYKADFTECCPADDKLALIP 203
 QY 181 KLDELREGRKASSAKORLKASLQKFGERAFKAWAVARLSORPPKAEFAEVSKLVTDLTK 240
 Db 204 KLDALKERILLSSAKERLKCSSQNFGERAVKAVSWARLSQKPKADFAEVSKLVTDLTK 263
 QY 241 VHTCCGGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCHIAEVENDEMPA 300
 Db 264 VHKCCGGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCHIAEVEDDLFS 323
 QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGMFLYEYARRHPYVSVLLRLAKTYETTLKXC 360
 Db 324 DLPALADFAEDKEICIKHYKDAKDVFLGTFLYEYARRHPYVSVLLRLAKTYETTLKXC 383
 QY 361 CAAADPHECYAKVDFEKLVEEPONLIKONCELFEOLEGEYKFNALLVRYTKKAPQVST 420
 Db 384 CAEDPPACYRTVDFQDFTPLVEEPKSLVKKNCDFEEVGEYDFQNALIVRYTKKAPQVST 443
 QY 421 PTLVEVSNLGVSKCKCKHPEAKMPCAEYDLSVLNQLCVLHEKTPVSDRVTKCTTES 480
 Db 444 PTLVEIGTLGVSRCCCLPESRLPCSENLALNRLCVLHEKTPVSKITKCTDS 503
 QY 481 LVNRPFCPSALEVDETYVYKFNATFTFHADICTLSEKQIKKQFALVELVXHKPKAT 540
 Db 504 LAERPCPSALELDEGVYVYKFNATFTFHADICTLSEKQIKKQFALVELVXHKPKAT 563
 QY 541 KEQLKAVMDDPFAFVEKCKADKCTCFABEGKLVAAQAAL 583
 Db 564 DEQLKTVNGDFGAFVEKCAENKEGCFSEBGPKLVAQAAL 606
 RESULT 6
 ALBU_BOVIN
 ID ALBU_BOVIN STANDARD; PRT; 607 AA.
 AC P02769; C02787;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serum albumin precursor (Allergen Bos d 6).
 GN ALB.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Holowachuk E.W., Stoltzenberg J.K., Reed R.G., Peters T. Jr.;
 RL Submitted (AUG-1991) to the EMBL/GenBank/DBSJ databases.

DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serum albumin precursor (Allergen Can f 3).
 GN ALB.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]_TaxID=9615;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Beagle; TISSUE=Liver;
 RA Hilger C.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=20148667; PubMed=10669848;
 RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H.,
 RA Valenta R., Spitzauer S.;
 RT "Escherichia coli expression and purification of recombinant dog
 RT albumin, a cross-reactive animal allergen.";
 RL J. Allergy Clin. Immunol. 105:279-285(2000).
 RN [3]
 RP SEQUENCE OF 25-48.
 RX MEDLINE=75011422; PubMed=4414013;
 RA Dixon J.W., Sarkar B.;
 RT "Isolation, amino acid sequence and copper(II)-binding properties of
 RT peptide (1-24) of dog serum albumin.";
 RL J. Biol. Chem. 249:5872-5877(1974).
 RN [4]
 RP SEQUENCE OF 25-38.
 RC TISSUE=Heart;
 RX MEDLINE=98163340; PubMed=9504812;
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
 RT dog heart proteins.";
 RL Electrophoresis 18:2795-2802(1997).
 RN [5]
 RP SEQUENCE OF 215-478 FROM N.A.
 RC TISSUE=Salivary Gland;
 RX MEDLINE=94201492; PubMed=7512102;
 RA Spitzauer S., Schweiger C., Sperr W.R., Pandjaitan B., Valent P.,
 RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;
 RT "Molecular characterization of dog albumin as a cross-reactive
 RT allergen.";
 RL J. Allergy Clin. Immunol. 93:614-627(1994).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- ALLERGEN: Causes an allergic reaction in human.
 CC -!- SIMILARITY: Belongs to the ALB/AFB/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; AJ133489; CAB564867.1; --
 DR EMBL; Y17737; CAB76841.1; --
 DR EMBL; S72946; AAB30434.1; --
 DR HSP; P02768; IE7B.
 DR HSC-2DPAGE; P49822;
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002466; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR

DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
 FT SIGNAL 1 18
 FT PROPEP 19 24 SERUM ALBUMIN.
 FT CHAIN 25 608
 FT DOMAIN 25 205 ALBUMIN 1.
 FT DOMAIN 212 397 ALBUMIN 2.
 FT DOMAIN 404 595 ALBUMIN 3.
 FT METAL 27 27 COPPER (BY SIMILARITY).
 FT DISULFID 77 86 BY SIMILARITY.
 FT DISULFID 99 115 BY SIMILARITY.
 FT DISULFID 114 125 BY SIMILARITY.
 FT DISULFID 148 193 BY SIMILARITY.
 FT DISULFID 192 201 BY SIMILARITY.
 FT DISULFID 224 270 BY SIMILARITY.
 FT DISULFID 269 277 BY SIMILARITY.
 FT DISULFID 289 303 BY SIMILARITY.
 FT DISULFID 302 313 BY SIMILARITY.
 FT DISULFID 340 385 BY SIMILARITY.
 FT DISULFID 384 393 BY SIMILARITY.
 FT DISULFID 416 462 BY SIMILARITY.
 FT DISULFID 461 472 BY SIMILARITY.
 FT DISULFID 485 501 BY SIMILARITY.
 FT DISULFID 500 511 BY SIMILARITY.
 FT DISULFID 538 583 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT CONFLICT 1 26 MKWTFISLFLFSSAYSRLVRREA -> MDT (IN
 FT REF. 2).
 FT CONFLICT 146 146 A -> R (IN REF. 2).
 FT CONFLICT 206 206 I -> T (IN REF. 2).
 FT CONFLICT 349 349 V -> A (IN REF. 2).
 FT CONFLICT 359 359 S -> A (IN REF. 1).
 FT CONFLICT 448 448 V -> VV (IN REF. 5).
 FT CONFLICT 474 474 D -> E (IN REF. 1).
 SQ SEQUENCE 608 AA; 3CF1C8FF7DD8FC06 CRC64;
 Query Match 82.6%; Score 2562; DB 1; Length 608;
 Best Local Similarity 79.8%; Pred. No. 4.2e-157;
 Matches 465; Conservative 57; Mismatches 61; Indels 0; Gaps 0;
 QY 1 DAHSEVAHFKDLGEENFKALVLIAPAQLQCCPDHVKLVNEYTERAKTCAVESAE 60
 DB 25 EAYKSEIAHRYNDLGEHFRGLVAVFSQYLOQCPDHDVYKLAKEVTEFAKAAESGA 84
 QY 61 NCDKSLHTLPGDKLCTVATRETYGEMADCCAKQEPERNECFQHKDDNPNLRLVRPEV 120
 DB 85 NCDKSLHTLPGDKLCTVATRETYGEMADCCAKQEPERNECFQHKDDNPNLRLVRPEV 144
 QY 121 DVMCTAFHDNEETFLKYLVEIARRHPYFYAPELLFFAKYKAAFTCCQAADKAACLLP 180
 DB 145 DALCAAFQDNEQLFLGKYLVEIARRHPYFYAPELLFYAQQYKGVFAECQQAADKAACLG 204
 QY 181 KLDELDEGKASSAKORLKCASLOKGERAFKAWARLSQRPKAFARVSKLVTDLTK 240
 DB 205 KIEALREKVLSSAKERFKCASLOKGERAFKAWARLSQRPKAFARVSKLVTDLTK 264
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLEKSHCHIAEVENDEMPA 300
 DB 265 VHKECCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPLEKSHCHIAEVENDEMPA 324
 QY 301 DLPSLAADFVSKDVKYKAEAKDVFGLMFLYVARHPDYSVVLALLRLAKTYETLEK 360
 DB 325 DLPSLAADFVSKDVKYKAEAKDVFGLMFLYVARHPDYSVVLALLRLAKTYETLEK 384
 QY 361 CAAADPHCEKAKVDFEKFPLVEEPQNLIKONCELFEOQKFEQNALLVRYTKKVPQVST 420
 DB 395 CATDDPTCYAKVLDEFKPLVDFEQNLVKTNCBELFEXGLGEYFGFNALLVRYTKKVPQVST 444
 QY 421 PTLVEVRNLGKVGSKCKKPEAKRMPCADYLSVNLQVLHKEKTPVSDRVTKCCTES 480
 DB 445 PTLVEVRNLGKVGSKCKKPEAKRMPCADYLSVNLQVLHKEKTPVSDRVTKCCTES 504
 QY 481 LVNRRPCFSALEVDVTVPKFEFNAETFTFHADICTLSEKERQIKKQATLVELVKHKPKAT 540

Query Match 100.0%; Score 3103; DB 1; Length 609;
 Best Local Similarity 100.0%; Pred. No. 9e-192;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEEFNFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
 DB 25 DAHKSEVAHRFKDLGEEFNFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 144
 QY 121 DVMCTAFHDNEETFLKKYLYEYIARHPHYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 145 DVMCTAFHDNEETFLKKYLYEYIARHPHYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 204
 QY 181 KLDELDEGKASSAKQRLKASLOKFGFRAPKAWAVARLSQFPKAEFAEVS KLVTDLTK 240
 DB 205 KLDELDEGKASSAKQRLKASLOKFGFRAPKAWAVARLSQFPKAEFAEVS KLVTDLTK 264
 QY 241 VHTTECHGDLLECCADRADLAKYICENQDISSSKLKSCCKEPLLEKSHCIAEVENDEMPA 300
 DB 265 VHTTECHGDLLECCADRADLAKYICENQDISSSKLKSCCKEPLLEKSHCIAEVENDEMPA 324
 QY 301 DLPSLAADFVSKDVKCKNYAKOVFLGMFLYEVARRHPDYSVVLLRLAKTVEYTTLEKC 360
 DB 325 DLPSLAADFVSKDVKCKNYAKOVFLGMFLYEVARRHPDYSVVLLRLAKTVEYTTLEKC 384
 QY 361 CAADAPHECYAKVDFEFPKPLVEEPPONLIKONCELFELQGEYKFNQALLVRYTKVPQVST 420
 DB 385 CAADAPHECYAKVDFEFPKPLVEEPPONLIKONCELFELQGEYKFNQALLVRYTKVPQVST 444
 QY 421 PTLVEVSRLNGLKVGSKCKHPKAEKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 445 PTLVEVSRLNGLKVGSKCKHPKAEKMPCAEDYLSVNLNQLCVLHEKTPVSDRVTKCCTES 504
 QY 481 LVNRRPCFSALSVDETYVPKFNFAETFFHADICTLSEKEROIKKQATLVELVHKPKAT 540
 DB 505 LVNRRPCFSALSVDETYVPKFNFAETFFHADICTLSEKEROIKKQATLVELVHKPKAT 564
 QY 541 KEQLKAVMDDFAAFEVKCKADDKTCFAEBGKKLVAASQAALGL 585
 DB 565 KEQLKAVMDDFAAFEVKCKADDKTCFAEBGKKLVAASQAALGL 609

RESULT 2
 ALBU_MACMU STANDARD; PRT; 600 AA.

ID ALBU_MACMU Q28522; AC
 01-NOV-1997 (Rel. 35, Created)
 01-NOV-1997 (Rel. 35, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor (Fragment).
 DB ALB.
 GN Macaca mulatta (Rhesus macaque).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Macaca.
 ON NCBI_TaxID=9544;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93211971; PubMed=8450152;
 RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
 RW Dulet J., Putnam F.W.;
 RT "cDNA and protein sequence of polymorphic macaque albumins that differ
 in bilirubin binding";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413 (1993).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 hormones, bilirubin and drugs. Its main function is the regulation
 of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.
 CC -!- SIMILARITY: Belongs to the ALB/APP/VDB family.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.
 CC EMBL; M90463; AAA36906.1; -;
 DR PIR; A47391; A47391.
 DR HSP; P02768; IE7B.
 DR InterPro; IPR000264; Serum albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum albumin; 1.
 DR SMART; SM00103; ALBUMIN_3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT NON_TER 1 1
 FT SIGNAL <1 10 BY SIMILARITY.
 FT PROPEP 11 16 BY SIMILARITY.
 FT CHAIN 17 600 SERUM ALBUMIN.
 FT DOMAIN 17 197 ALBUMIN 1.
 FT DOMAIN 204 289 ALBUMIN 2.
 FT DOMAIN 396 597 ALBUMIN 3.
 FT METAL 19 19 COPPER (BY SIMILARITY).
 FT BINDING 256 256 BILIRUBIN (POTENTIAL).
 FT DISULFID 69 78 BY SIMILARITY.
 FT DISULFID 91 107 BY SIMILARITY.
 FT DISULFID 106 117 BY SIMILARITY.
 FT DISULFID 140 185 BY SIMILARITY.
 FT DISULFID 184 193 BY SIMILARITY.
 FT DISULFID 216 262 BY SIMILARITY.
 FT DISULFID 261 269 BY SIMILARITY.
 FT DISULFID 281 295 BY SIMILARITY.
 FT DISULFID 294 305 BY SIMILARITY.
 FT DISULFID 332 377 BY SIMILARITY.
 FT DISULFID 376 385 BY SIMILARITY.
 FT DISULFID 408 454 BY SIMILARITY.
 FT DISULFID 453 464 BY SIMILARITY.
 FT DISULFID 477 493 BY SIMILARITY.
 FT DISULFID 492 503 BY SIMILARITY.
 FT DISULFID 530 575 BY SIMILARITY.
 FT DISULFID 574 583 BY SIMILARITY.
 SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match 94.8%; Score 2942; DB 1; Length 600;
 Best Local Similarity 93.5%; Pred. No. 1.8e-181;
 Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEEFNFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
 DB 17 DTHKSEVAHRFKDLGEEFNFKALVLIAPAQYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 76
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 120
 DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPNLRLVRPEV 136
 QY 121 DVMCTAFHDNEETFLKKYLYEYIARHPHYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 DB 137 DVMCTAFHDNEETFLKKYLYEYIARHPHYFAPPELLFFAKRYKAAFTCCQAADKAACLLP 196
 QY 181 KLDELDEGKASSAKQRLKASLOKFGFRAPKAWAVARLSQFPKAEFAEVS KLVTDLTK 240
 DB 197 KLDELDEGKASSAKQRLKASLOKFGFRAPKAWAVARLSQFPKAEFAEVS KLVTDLTK 256
 QY 241 VHTTECHGDLLECCADRADLAKYICENQDISSSKLKSCCKEPLLEKSHCIAEVENDEMPA 300
 DB 257 VHTTECHGDLLECCADRADLAKYICENQDISSSKLKSCCKEPLLEKSHCIAEVENDEMPA 316

RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci F., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting R.W., Touchman A., Young A.C., Sherchenko Y., Bouffard G.G.,
RA Blakesley M., Madan A., Young A.C., Sherchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RP SEQUENCE OF 25-609.
RX Brown J.R., Shockley P., Behrens P.O.;
RL (In) Bing D.H. (eds.);
RT The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RP SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Menaya J., Farrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=24193129;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RP SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RP SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994.";
RL FEBS Lett. 15:1459-1465(1994).
RN [14]
RP DISULFIDE BONDS.
RA Saber M.A., Stockbauer P., Moravsek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15]
RP BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin.";
RL Biochem. J. 171:453-459(1978).

RN [16]
RP VARIANT CANTERBURY ASN-337.
RX MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Lys-->Asn). A point mutation in the second
RT domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RP VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=89068522; PubMed=3479777;
RA Takahashi N., Takanashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
RT Amerindian and Japanese populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18]
RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huss K., Ishioke N., Satoh C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese alloalbumins.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19]
RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20]
RP DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Wyles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala-->Thr): a glycoprotein variant of
RT human serum albumin whose precursor has an aberrant signal peptidase
RT cleavage site.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21]
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
RT Italy.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22]
RP VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
RT carboxyl-terminal variants of human serum albumin.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23]
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RA MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Anaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN [24]
RP VARIANT CASEBOOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
RT albumin: albumin Casebook (494 Asp-->Asn).";
RL Biochim. Biophys. Acta 1097:49-54(1991).
RN [25]
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two alloalbumins with identical electrophoretic mobility are produced

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:25:34 ; Search time 42.6731 Seconds
(without alignments)
713.823 Million cell updates/sec

Title: US-09-832-929-18
Perfect score: 3103
Sequence: 1 DAHSEVAHRFDLGEENFK.....TCFAEEGKLVASQAALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	609	1	ALBU HUMAN
2	2942	94.8	600	1	P02768 macaca sapien
3	2620	84.4	608	1	P49064 felis silve
4	2562	82.6	608	1	P49822 canis faml
5	2475.5	79.8	607	1	P35747 equus cabal
6	2450.5	79.0	607	1	P02769 bos taurus
7	2446	78.8	608	1	P49065 oryctolagus
8	2432.5	78.4	607	1	P14539 ovies aries
9	2426	78.2	608	1	P02770 rattus norv
10	2411.5	77.7	605	1	P08835 sus scrofa
11	2387	76.9	609	1	O35890 meriones un
12	2378	76.6	608	1	P07724 mus muscull
13	1557.5	50.2	615	1	P19121 gallus gall
14	1253.5	40.4	609	1	Q28789 pan troglod
15	1249.5	40.3	609	1	P02771 homo sapien
16	1242.5	40.0	609	1	P28050 gorilla gor
17	1203	38.8	607	1	P14872 xenopus lae
18	1200	38.7	609	1	P49066 equus cabal
19	1164.5	37.5	606	1	P08759 xenopus lae
20	1084	34.9	605	1	P02772 mus muscull
21	1067	34.4	611	1	P02773 rattus norv
22	1055	34.0	599	1	P43652 homo sapien
23	944	30.4	611	1	O89020 mus sapien
24	928	29.9	608	1	P16853 rattus norv
25	747.5	24.1	608	1	P21848 salmo salar
26	742.5	23.9	608	1	Q03156 salmo salar
27	699	22.5	382	1	P21847 rana catesb
28	440.5	14.2	1423	1	Q91374 petromyzon
29	386	12.4	476	1	P04274 homo sapien
30	381	12.3	476	1	P04276 rattus norv
31	378	12.2	476	1	P23789 oryctolagus
32	372	12.0	476	1	P21814 mus muscull
33	151.5	4.9	1605	1	Q99P15 mus muscull

ALIGNMENTS

RESULT 1

ALBU_HUMAN	144.5	4.7	8797	1	SNE1_HUMAN	O8nf91	
AC	P02768; O95574; Q13140; Q9P157; Q9PLI7; Q9UHS3; Q9UJZ0;	1410	1	RRB1_HUMAN	Q9pde9		
DT	21-JUL-1986 (Rel. 01, Created)	4.3	1391	1	MST2_DROHY	Q08696	
DT	01-APR-1990 (Rel. 14, Last sequence update)	4.3	2230	1	GOA4_HUMAN	Q13439	
DT	15-MAR-2004 (Rel. 43, Last annotation update)	4.2	1972	1	MYH8_HUMAN	P35749	
DE	Serum albumin precursor.	129	4.2	3210	1	CENF_HUMAN	P49454
GN	ALB.	128	4.1	1005	1	RA50_METJA	O58718
OS	Homo sapiens (Human).	126.5	4.1	1972	1	MYH8_RABIT	P35748
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	126	4.1	1189	1	SMC2_CHICK	Q90988
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	126	4.1	3259	1	GOB1_HUMAN	Q14789
OX	NCBI_TaxID=9606;	125	4.0	3038	1	TRIO_HUMAN	O75962
RN	[1]	124.5	4.0	1790	1	USO1_YEAST	P25386
RP	SEQUENCE FROM N.A.						
RX	MEDLINE=86196112; PubMed=3009475;						
RA	Minghetti P.P., Ruffner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,						
RA	Beattie W.G., Dugaiczky A.;						
RT	"Molecular structure of the human albumin gene is revealed by						
RT	nucleotide sequence within q11-22 of chromosome 4.";						
RL	J. Biol. Chem. 261:6747-6757(1986).						
RN	[2]						
RP	SEQUENCE FROM N.A., AND VARIANT LYS-420.						
RX	MEDLINE=82081882; PubMed=6171778;						
RA	Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,						
RA	Najarian R.C., Seeburg P.H., Wion K.L.;						
RT	"The sequence of human serum albumin cDNA and its expression in E.						
RT	coli.";						
RL	Nucleic Acids Res. 9:6103-6114(1981).						
RN	[3]						
RP	SEQUENCE FROM N.A., AND VARIANT GLY-121.						
RX	MEDLINE=82105994; PubMed=6275391;						
RA	Dugaiczky A., Law S.W., Dennison O.E.;						
RT	"Nucleotide sequence and the encoded amino acids of human serum						
RT	albumin mRNA.";						
RL	Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).						
RN	[4]						
RP	SEQUENCE FROM N.A.						
RC	TISSUE=Liver;						
RA	Yang S., Zhang R.A., Qi Z.W., Yuan Z.Y.;						
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.						
RN	[5]						
RP	SEQUENCE FROM N.A. (PRO00903/PRO1708/PRO2044/PRO2619/PRO2675).						
RC	TISSUE=Fetal liver;						
RA	Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y.,						
RA	Xu W., Gao F., Liu M., He F., Zhang Y., Ouyang S., Luo L.;						
RT	"Functional prediction of the coding sequences of 121 new genes						
RT	deduced by analysis of cDNA clones from human fetal liver.";						
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.						
RN	[6]						
RP	SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.						
RA	Huang M.C., Wu H.T.;						
RT	"The cDNA sequences of human serum albumin.";						
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.						
RN	[7]						
RP	SEQUENCE FROM N.A.						

```
Db 141 VPEPVTSCAEYEDRETFMNFYIARRHPFLYAPTILLWAARYDKLIIPSCCKAENAVE 200
Qy 177 CLLPKLDLRDEGKASSAKRLKASLOKFGERAFAKAWARLSORPPKAFASVSKLVT 236
Db 201 CFQTKAAVTKELRESSLLNQHACAVMKNFGTRTFOAITVTKLSOKFTKVNFTSLOKLVL 260
Qy 237 DLTQVHTECHCHDLLECCADRDADLAKYICENQDSISSKLKCECKPLLEKSHCIAEVEND 296
Db 261 DVAVHVEHCRCRDVLDCLQDGKINSYICSQDQTLNSKITECKLTLTLERGQCIIHAEND 320
Qy 297 ENPADLPSLAADPVESKDVCKVQYAAKOVFLGMFLYEVARRHPDVSVLLRLAKTYETT 356
Db 321 EXPEGLSPNLNRLFDGDRNPFQSSGKNIIFLASFVHEYSRRHPQAVSVILRVAKGYOEL 380
Qy 357 LEKCCAAADPHCYAKVDFEFKPLVEEPONLIKONCELFEOIGYKFNALLVRYTKVP 416
Db 381 LEKCFQTEPNLSCQDGHEELQYIOESQALAKRSQGLFQKLGYYLQNAFLVATKKAP 440
Qy 417 QVSTPTLVEVSRNLKGVGKCKKHPKAGMPCABDYLSVVLNQLCVLHEKTPVSDRVTKC 476
Db 441 QLTSSLMATRKMAATAATCCQLSEDKLLACGEGAADIIGHLCIRHEMTVPNPGVGQC 500
Qy 477 CTESLVNRRPCSALEVDVTVYKPFNAETFTFHADICTLSEKROIKKQTALVELVKHK 536
Db 501 CHSSVANRRPCFSSLVVDVTVYPPAFSDDKFPHKDLCOAQVQALQTMKQEFNLINVKQK 560
Qy 537 PRATKEQLKAVMDDFAAFVEKCKCKADDETCFAEBGKKLVAASQAALGL 585
Db 561 PQITEQLEVTIADFSGILEKCCQGOEQVCFABEGCKLISKTRTALGV 609
```

RESULT 15

```
ABXL72
74K albumin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: B41682; S02693; A05288
R:Moskaitis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.
A:Reference number: A41682; MUID:89313788; PMID:2747653
A:Accession: B41682
A:Molecule type: mRNA
A:Residues: 3-607 <SOS>
A:Cross-references: GB:M21442; NID:G213930; PIDN:AAA49637.1; PID:G213931
R:Schorpp, M.; Doebeiling, U.; Wagner, U.; Rytzel, G.U.
J. Mol. Biol. 199, 83-93, 1988
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Deletion of the 5'-flanking region of the 5'-proximal exon results in a nonfunctional gene.
A:Reference number: S02692; MUID:88172470; PMID:2451026
A:Accession: S02693
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-48 <SCH>
A:Cross-references: EMBL:Z26826
R:Wolffe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, J. Eur. J. Biochem. 146, 489-496, 1985
A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabilization of the 74-kDa albumin mRNA.
A:Reference number: A05288; MUID:85126974; PMID:3971963
A:Accession: A05288
A:Molecule type: mRNA
A:Residues: 459-502, 'L', 504-557 <WOL>
A:Cross-references: GB:M28276
A:Note: the authors translated the codon TAT for residue 63 as Thr
C:Comment: Serum albumin is synthesized in the liver as prealbumin. It binds copper, mones (weak bonds with these hormones promote their transfer across the membranes), thyroglobulin, and various other ligands.
C:Genetics: 27/1
A:Introns: 27/1
C:Superfamily: serum albumin; serum albumin repeat homolog
C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
```

```
F:25-607/Product: 74K serum albumin #status predicted <MAT>
F:32-201/Domain: serum albumin repeat homolog <SA1>
F:220-393/Domain: serum albumin repeat homolog <SA2>
F:412-591/Domain: serum albumin repeat homolog <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-88/101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392
F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 38.8%; Score 1205; DB 1; Length 607;
Best Local Similarity 39.3%; Pred. No. 3.3e-72;
Matches 227; Conservative 108; Mismatches 239; Indels 4; Gaps 2;
```

```
Qy 3 HKSEVAHRFDKLGSENEFKALVLIAPQYLOQCPFFEDHVKLVNEVTEFAKTCVADGEAENC 62
Db 30 HHKHIAADVYALTERTFKGTJLTAIVSQNLQKSELSKLVNEINDFAKSINDKTIPE-C 88
Qy 63 DKSLHTLFGKLCITVATLRETYGEMADCCAKQBPENNECFLOKHQDNPNLPLRVREVDV 122
Db 89 EKPVGTLFFDKLCADPAVGVNYSKECCAKQDPERAQCFKAHRDHEHT--SIKPEPEE 145
Qy 123 MCTAFHDNETFLKVLVEIARRHPYFAPPELLFPAKRYKAAFTCCQAADKACLLPKL 182
Db 146 TCCLKKEHPDDLLSAFTHEARNHPDLYPPAVLALTQVHKLAHCEDEEKEKCFSEKM 205
Qy 183 DELRDECKASSAKQRLKCASLQKFGERAFAKAWARLSORPPKAFASVSKLVTDLTKVH 242
Db 206 KQLMKQSHSIEDQHHFCWILDNFPKVLKALNLAARVSHYPAEFKLAHNFTEEVTHFI 265
Qy 243 TECCHGDLLECADRDADLAKYICENQDSISSKLKCECKPLLEKSHCIAEVENDMPADL 302
Db 266 KDCCHDDMFECMTERTLEHTCQHKDELSSKLEKCCNIPLLETTYCIVTLENDVPAEL 325
Qy 303 PSLAADPVESKDVCKVQYAAKOVFLGMFLYEVARRHPDVSVLLRLAKTYETTLEKCCA 362
Db 326 SQPITEFTEDPHVCEKYAENNEVFLGRLHAVSKKQELSEQFLQSAKESYLKCKCK 385
Qy 363 AADPHCYAKVDFEFKPLVEEPONLIKONCELFEOIGYKFNALLVRYTKVPQVSTPT 422
Db 386 TDNPECYKQDADRFMNEAKERFAYLKQNCIDLHEHGEYLFENELLRYTKQMPQVSDET 445
Qy 423 LVEVSRNLKGVGKCKKHPKAGMPCABDYLSVVLNQLCVLHEKTPVSDRVTKCTESLV 482
Db 446 LIGAHQADIGEHCACAVPENQRMPCAGDGLTILIGKRCERQKTFINNVHVAICCTDSYS 505
Qy 483 NRRPCFSALEVDVTVYKPFNAETFTFHADICTLSEKROIKKQTALVELVKHKPKATKE 542
Db 506 GMRSCFTALGDEDDYVPPVTDDTFHDDKICTANDKEKHQIKQKFLVKLVKVPKLEKN 565
Qy 543 QLKAVMDDFAAFVEKCKCKADDETCFAEBGKKLVAASQ 580
Db 566 HIDECSAEFLKQVQKCCCTADEHQPCFDTKEPVLIEHCQ 603
```

Search completed: April 19, 2004, 12:02:19
Job time : 73.3019 secs

A;Residues: 19-45;63-97;102-107;122-184;187-249;255-489;507-609 <TEC>
R;Yachnin, S.; Heu, R.; Heinrikson, R.L.; Miller, J.B.
Biochim. Biophys. Acta 493, 418-428, 1977
A;Title: Studies on human alpha-fetoprotein. Isolation and characterization of monomeric
A;Reference number: A90624; MUID:77242506; PMID:70228
A;Accession: A90624
A;Molecule type: protein
A;Residues: 'S', 20-22, 'S', 24-35 <YAC>
A;Note: dimeric and trimeric forms have been found in addition to the monomeric form
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 37, 3663-3667, 1977
A;Title: Comparative chemical structure of human alpha-fetoproteins from fetal serum and
A;Reference number: A90757; MUID:78001760; PMID:71198
A;Accession: A90757
A;Molecule type: protein
A;Residues: 'S', 20-30, 'A', 32-37, 'A' <AOY>
R;Ruohlahti, E.; Piikko, H.; Vaheri, A.; Seppala, M.; Virolainen, M.; Konttinen, A.
Johns Hopkins Med. J. Suppl. 3, 249-255, 1974
A;Title: 20. Alpha fetoprotein: structure and expression in man and inbred mouse strains
A;Reference number: A93042; MUID:75018719; PMID:4138095
A;Accession: A93042
A;Molecule type: protein
A;Residues: 'S', 20-24, 'Q', 26-30, 'A', 32-35, 'E', 37-39 <RUO>
R;Sakai, M.; Morinaga, T.; Urano, Y.; Watanabe, K.; Wegmann, T.G.; Tamaoki, T.
J. Biol. Chem. 260, 5055-5060, 1985
A;Title: The human alpha-fetoprotein gene. Sequence organization and the 5' flanking reg
A;Reference number: A92520; MUID:85182629; PMID:2580830
A;Contents: annotation; gene, exons and introns
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.
A;Reference number: A90758; MUID:79001617; PMID:80265
A;Contents: annotation; metal binding
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 39, 3571-3574, 1979
A;Title: alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding abil
A;Reference number: A90759; MUID:80001710; PMID:89900
A;Contents: annotation; bilirubin binding
C;Comment: AFP is synthesized by the fetal liver and yolk sac. It occurs in the plasma o
o trace amounts after birth. The serum level in adults is usually less than 40 ng/ml. AF
C;Comment: Human AFP binds copper, nickel, and fatty acids as well as, and the bilirubin
Properties.
C;Genetics:
A;Gene: GDB:AFP
A;Cross-references: GDB:119660; OMIM:104150
A;Map position: 4q11-4q13
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-609/Product: alpha-fetoprotein #status experimental <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;22/Binding site: copper (His) #status experimental
F;99-114, 113-124, 148-193, 192-201, 224-270, 269-277, 289-303, 302-313, 384-393, 416-462, 461-472
F;249/Binding site: bilirubin (Lys) #status predicted
F;251/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 40.3%; Score 1249.5; DB 1; Length 609;
Best Local Similarity 39.9%; Pred. No. 3.7e-75;
Matches 235; Conservative 116; Mismatches 231; Indels 7; Gaps 3;
QY 3 HKSE-----VAHRRFDLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADE 57
DB 22 HRNEYGIASILDYSQCTAIEISLADLATAFFAQFVQEAITYKEVSKVNDALTALEKPTGDE 81
QY 58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERCEFLQHKDNP-NLPLRV 116
DB 82 QSSGLENQLPAFLBELCHEKEILEKYGH-SDCCSQSSEGRHNCFLAHHKPTASIPLFQ 140
QY 117 RPEVDVMTAFHDNETFLKYLVIARRHPYFVAPELLFFAKRYKAAFTCCQAADKAA 176

Db 141 VBPVTSCEAYEDRETFMKNFYIETARRHPFLYATFILLWAARYDKIIPSCCKAENAVE 200
QY 177 CLFLPKDLDELDEGKASSAKQRLKASLQKFGEPAPKAWAVARLSQFFPKAFAEVSKLVT 236
Db 201 CFQTKAATVTKELRESSLLNQHACAVMKNFEGTTFQAITVTKLQSKFTKVNFTIQLVL 260
QY 237 DLTQVHTCECHGDLLECADRADLAKYICENQDISSKLKECCPKLLEKSHCIAEVEND 296
Db 261 DVAIVHHCRCRGDVLDCQGEKIMSYICQQDTLSNKTCCCKLTLEGGQCIHAEND 320
QY 297 EMPADLPSLAADFYESKDVCKVNAEAKDVLGMLFLYEAARRHPDYSVLLRLAKYETT 356
Db 321 EKPEGLSPNLNRFLGDRDFNQFSSGEKNIFLASFVHEYSRRHPQLAVSVILRVAKGYOEL 380
QY 357 LEKCCAAADPHECAKVFDFEFKPLVEEPQNLIKQCELFECGLGEYKFTQNALLYRTKVP 416
Db 381 LEKCFQTEPLECDKGEELQKYIQESQALAKRSCGLFQKLGLEYLQNAFLVAYTKAP 440
QY 417 QVSTPTLVEVSRLNGKVGSKCCKHPEAKRMPCAEDYLSVLNLQCVLHEKTPVSDRVTKC 476
Db 441 QLTSSSELMAITRKMAATAATCCQLSEDKLLACGSGAADIIGHLCIRHEMTVPNPGVGQC 500
QY 477 CTESLVNRRPCFSALEVDVTVPEFNAETTFHADICTLSEKERQIKKQATLVELVKHK 536
Db 501 CTSSYANRRPCFSLVVDVTVPPAFDDKPIFFHKDJCOAGGVALQTMKQEFNLNLVKQK 560
QY 537 PKATKEQLKAVMDFFAAVFEKCCCKADDKTCFAEKGKLVAAQAALGL 585
Db 561 PQITEQLEAVIADFSGLLEKCCQGEVCFAEKGKLVKTRAAALGV 609
RESULT 14
PFCO
alpha-fetoprotein precursor - gorilla
C;Species: Gorilla gorilla (gorilla)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C;Accession: A37970
R;Ryan, S.C.; Zielinski, R.; Dugaiczky, A.
Genomics 9, 60-72, 1991
A;Title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primate
A;Reference number: A37970; MUID:91169517; PMID:1706310
A;Accession: A37970
A;Molecule type: DNA
A;Residues: 1-609 <RYA>
A;Cross-references: GS:M38272; NID:9817963; PIDN:AAA73520.1; PID:G177041
C;Genetics:
A;Map position: 4q11-12
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 51
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-609/Product: alpha-fetoprotein #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;22/Binding site: copper (His) #status predicted
F;99-114, 113-124, 148-193, 192-201, 224-270, 269-277, 289-303, 302-313, 384-393, 416-462, 461-4
F;249/Binding site: bilirubin (Lys) #status predicted
F;251/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 40.0%; Score 1242.5; DB 1; Length 609;
Best Local Similarity 39.6%; Pred. No. 1.1e-74;
Matches 233; Conservative 117; Mismatches 232; Indels 7; Gaps 3;
QY 3 HKSE-----VAHRRFDLGEENFKALVIAFAQYLQCCPFEDHVKLVNEVTEFAKTCVADE 57
DB 22 HRNEYGIASILDYSQCTAIEISLADLATAFFAQFVQEAITYKEVSKVNDALTALEKPTGDE 81
QY 58 SAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERCEFLQHKDNP-NLPLRV 116
Db 82 QSSGLENQLPAFLBELCHEKEILEKYG-LSDCSQSSEGRHNCFLAHHKPTASIPLFQ 140
QY 117 RPEVDVMTAFHDNETFLKYLVIARRHPYFVAPELLFFAKRYKAAFTCCQAADKAA 176


```

Db      330 LPSLVEKTYEDKEVCKSPFAGHDAMFASVFVYSRRHDFEFTQILIMRIAKGYESLLEKCC 389
QY      362 AAADPHECYAKVDFEFKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKVPQVSTP 421
Db      390 KTDNPAECYANAQEQNLQHIKETQDVVKTNCOLLHDHGEADPLKSLIRYTKQMPQVPTD 449
QY      422 TLVEVSRNLGVGSKCKHPKAPKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKCTESL 481
Db      450 LLELTKGKMTTIGTKCCQJGEDRRMACSEGYLSIVIHDTCKRQETTPINDNVQCCSOLY 509
QY      482 VNRRPCFSALEYDETYVPKFEAFETTFHADICTLSEKQRIKKQTALVELVHKPKATK 541
Db      510 ANRPPCFTAMGVDTKYVPPFPNDFMFSDEKLCSAPASEREVGQMKLLINLIKRPQMT 569
QY      542 EQLKAVMDPFAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
Db      570 EQIKTIADGFTAMVCKCKQSDINTCFGEEGANLVQSRATLGI 613

RESULT 13
JC4258
alpha-fetoprotein precursor - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 27-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 20-Aug-1999
C;Accession: JC4258
R;Nishio, H.; Gibbs, P.E.M.; Minghetti, P.P.; Zielinski, R.; Dugaiczky, A.
Gene 162, 213-220, 1995
A;Title: The chimpanzee alpha-fetoprotein-encoding gene shows structural similarity to
A;Reference number: JC4258; MUID:96032345; PMID:7557431
A;Accession: JC4258
A;Molecule type: DNA
A;Residues: 1-609 <NIS>
A;Cross-references: GB:U21916; NID:G841311; PIDN:AAA91641.1; PID:G841312
C;Comment: This protein is a plasma protein produced in the fetal and neonatal liver and
o similar properties and structure.
C;Genetics:
A;Gene: afp
A;Map position: 3p
A;Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3; 551
A;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: glycoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-609/Product: alpha-fetoprotein #status predicted <NAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;42,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.4%; Score 1253.5; DB 2; Length 609;
Best Local Similarity 40.1%; Pred. No. 2e-75;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

QY      3 HKSE-----VAHFRKDLGEENFKALVLIAPAYLOQCPPEFHVKLVNVEYTFKACTVADE 57
Db      22 HRNEYGIASILDSVQCTAEINLTDLATIFFAQVQBATYKSVKVKDALTAIEKPTGDE 81
QY      58 SAENCUKSLHTLFGDKLCTVATIRETYGEMADCAQEBERNECFLOHKDDNP-NLPLRV 116
Db      82 QSAGCUENQAPFLEELCREKEILEKYGH-SDCCSQSEGRHNCFLAHKKPTPASIPFFQ 140
QY      117 RPEVDVMTAFHONETFLKLYEIAARRHPFYAPPELLFPAKRYKAAFTCCQAAADKAA 176
Db      141 VPEPVTSCAYEEDREFMKNFIVEIARRHPFLYAPTILLWAARYDKIIPSCCKAENAVE 200
QY      177 CLIPKLDLDEKQKASSAKORLKASLQKFGERAPKAWAVARLSQRFPKAEFAEVSKLVT 236
Db      201 CFQTKATVTKELRESSLLNHQACAVMKNFGTFTQAITVTKLSQKFTKVNFEIQKLV 260
QY      237 DLTQVHTECGHDLLECADDRADLAKYICENQDISSKLKECCCKPLLEKSKHICIAEVND 296
Db      261 DVAVHVEHCRRGVLDCLQGEKIMSYICSQQTLSNKITECCCKLTILRGQCIIHAEND 320
QY      297 EMPADLPSLAADFVESKDVCKNVANAKDVFGLMFLVEYARRHPDYSVLLLLRLAKTYETT 356

```

```

Db      321 EKTEGSLSPNLRFLGDRDNQPSSEGEKNIFLASFVHEYSERHQPOLAVSVILRAKGYQEL 380
QY      357 LEKCCAAADPHECYAKVDFEFKPLVEEPONLIKONCELFQOLGEYKFNALLVRYTKKVP 416
Db      381 LEKCFOTENPLECQDGEELQYIQESQALAKRSCGLFKLGEYYLQNAFLVAYTKAP 440
QY      417 QVSTPTLVEVSRNLGVGSKCKHPKAPKMPKCAEDYLSVVLNQLCVLHEKTPVSDRVTKC 476
Db      441 QLTSSELMATRMAATAATCCQLSDKLLACGEAGADIIIGHLCITRHEHTTTPNPGVGQC 500
QY      477 CTSLVNRRPCFSALEYDETYVPKFEAFETTFHADICTLSEKQRIKKQTALVELVHK 536
Db      501 CTSSYANRRPCFSSLVVDETYVPFASDDKFIHKDLCOAQGVQALQTMKQEFNLINLVKQ 560
QY      537 PKATKEQLKAVMDPFAFVEKCKKADDKETCFABEGKLVAAASQAALGL 585
Db      561 PQITEGLEAVIADFSGLLEKCCQGEQVCFABEGKLVKSTRALLGV 609

RESULT 13
PPHU
alpha-fetoprotein precursor [validated] - human
A;Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin
C;Species: Homo sapiens (man)
C;Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 08-Dec-2000
C;Accession: A26624; S37655; A93961; A91497; A23699; A61480; A90624; A90757; A93042; A
R;Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczky, A.
Biochemistry 26, 1333-1343, 1987
A;Title: Structure, polymorphism, and novel repeated DNA elements revealed by a complet
A;Reference number: A26624; MUID:87195438; PMID:2436661
A;Accession: A26624
A;Molecule type: DNA
A;Residues: 1-609 <GIB>
A;Cross-references: GB:M6110; NID:G773678; PIDN:AAB58754.1; PID:G178236
R;McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilghman, S.; Krumlat
Hum. Mol. Genet. 2, 379-384, 1993
A;Title: A G->A substitution in an HNF I binding site in the human alpha-fetoprotein ge
A;Reference number: S37655; MUID:93279385; PMID:7684942
A;Accession: S37655
A;Molecule type: DNA
A;Residues: 1-28 <MCV>
A;Cross-references: EMBL:Z19532; NID:G28527; PIDN:CAA79592.1; PID:G28528
A;Note: the authors translated the codon TAT for residue 26 as Thr
R;Morinaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.
Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983
A;Title: Primary structures of human alpha-fetoprotein and its mRNA.
A;Reference number: A93961; MUID:83273664; PMID:6192439
A;Accession: A93961
A;Molecule type: mRNA
A;Residues: 1-609 <MOR>
A;Cross-references: GB:J00077; NID:G311348; PIDN:CAA24758.1; PID:G31351
R;Beattie, W.G.; Dugaiczky, A.
Gene 20, 415-422, 1982
A;Title: Structure and evolution of human alpha-fetoprotein deduced from partial sequen
A;Reference number: A91497; MUID:83158778; PMID:6187626
A;Accession: A91497
A;Molecule type: mRNA
A;Residues: 423-556 <BEA>
A;Cross-references: GB:J00076
R;Pucici, P.; Scilliano, R.; Malorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Terrar
Biochemistry 30, 5061-5066, 1991
A;Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.
A;Reference number: A23699; MUID:91242409; PMID:1709810
A;Accession: A23699
A;Molecule type: protein
A;Residues: 19-45;160-97;102-107;122-184;187-249;255-489;507-609 <PUC>
R;Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.
J. Nucl. Med. Allied Sci. 34, 213-216, 1990
A;Title: Characterization of in vitro expressed human alpha-fetoprotein as highly reprc
A;Reference number: A61480; MUID:91225826; PMID:1709209
A;Accession: A61480
A;Molecule type: protein

```



```
QY 422 TLVEVSRNLGKVGSKCKRHPKARMPCAEDYLVSVLNQLCVLHKTPVSDRVTKCTESL 481
DB 447 TLVEAARSLSRGVTHCCALPEKRLPCVEDYLSAILNRVCLLHKTVPSEQVTKCCSGSL 506
QY 482 VNRPCFSALEVDVETVVPKEFNAETFFHADICTLSEKERQIKKOTALVELVHKPKATK 541
DB 507 VERRPCFSALFVDVETVVPKEFNAETFFHADICTLSEKERQIKKOTALVELVHKPKATK 566
QY 542 EQLKAVMDPFAAFVEKCCCKADKCTCFABEGKKLVAASQAAL 583
DB 567 EQLKAVMDPFAAFVEKCCCKADKCTCFABEGKKLVAASQAAL 608
RESULT 10
A05139
serum albumin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 05-Jun-1987 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C:Accession: A05139; I48638
R:Minigetti, P.P.; Law, S.W.; Dugaiczky, A.
Mol. Biol. Evol. 2, 347-358, 1985
A:Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudog
A:Reference number: A93055; MUID:88216123; PMID:2452956
A:Accession: A05139
A:Molecule type: mRNA
A:Residues: 1-418 <MIN>
A:Cross-references: GB:M16111; NID:G191764; PIDN:AAA37190.1; PID:G191765
R:Boccacchio, C.; Deschatrete, J.; Meunier-Rotival, M.
Gene 88, 181-186, 1990
A:Title: Empty and occupied insertion site of the truncated LINE-1 repeat located in the
A:Reference number: I48638; MUID:90269606; PMID:1971802
A:Accession: I48638
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 379-453 <BOC>
A:Cross-references: EMBL:X13060; NID:G52939; PIDN:CAA31458.1; PID:G899334
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-104/Domain: serum albumin repeat homology (fragment) <SA1>
F:123-296/Domain: serum albumin repeat homology <SA2>
F:335-453/Domain: serum albumin repeat homology (fragment) <SA3>
Query Match 50.0%; Score 1861; DB 2; Length 453;
Best Local Similarity 72.2%; Pred. No. 9.4e-116;
Matches 327; Conservative 64; Mismatches 62; Indels 0; Gaps 0;
QY 75 CTATLRETYGEMADCCAKOBERNECFLOHKDDNPNLRLVPRVDMCTAFHNEETP 134
DB 1 CAIENLRENYGELADCTCKOBERNECFLOHKDDNPNLRLVPRVDMCTAFHNEETP 60
QY 135 LKCYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAAALLPKLDELREKASSA 194
DB 61 MGHYLEVARRHPYFYAPPELLFYAEQYNEILTOCCABADKESCLTPKLDGVKEKALVSV 120
QY 195 KORLKCSLOKFGERAFAKAVARLSORFPKAEVSKLAVTDLTGVHTECCGGDLLECA 254
DB 121 RQRVKCSMQKFGERAFAKAVARLSOTFFNADPAETIKLATDTKVKNECCGGDLLECA 180
QY 255 DDRADLAKYICENDSISSKLKECCFKPLLEKSHCIAEVENDEMPADLPISLADPVESKD 314
DB 181 DDRAELAKYMCENQATISSKLTCCDKPLLKKAHCLSEVHDTPADLPISLADPVESKD 240
QY 315 VCKRYAEAKDVLGFLMFLYEYARRHPDYSVLLLLAKTYETLEKCCAAADPHECYAKVF 374
DB 241 VCKRYAEAKDVLGFLMFLYEYARRHPDYSVLLLLAKTYETLEKCCAAADPHECYAKVF 300
QY 375 DEFKPLVEEPONLAKONCELPQOIGYKFNQNALVRYTKVPQVSTPTLVEVSRNLGKVG 434
DB 301 AEFQPLVEEPONLAKONCELPQOIGYKFNQNALVRYTKVPQVSTPTLVEVSRNLGKVG 360
QY 435 SKCCFKHPKARMPCAEDYLVSVLNQLCVLHKTPVSDRVTKCTESLWNRPCFSALEVD 494
DB 361 TKCCTPLDQLPCVEDYLSAILNRVCLLHKTVPSEHVTKCCSGSLVRRPCFSALTVD 420
```

```
QY 495 ETVVPKEFNAETFFHADICTLSEKERQIKKOT 527
DB 421 ETVVPKEFNAETFFHADICTLSEKERQIKKOT 453
```

RESULT 11

```
ABCHS
serum albumin precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: S15571; A05078; A13451
R:Cassady, A.I.; Salkild, C.K.; Baverstock, P.; Wallace, J.C.
submitted to the EMBL Data Library, July 1991
A:Reference number: S15571
A:Accession: S15571
A:Molecule type: mRNA
A:Residues: 1-615 <CAS>
A:Cross-references: EMBL:X60688; NID:G63747; PIDN:CAA43098.1; PID:G63748
R:Hache, R.J.G.; Wiskocil, R.; Vasa, M.; Roy, P.C.K.; Deeley, R.G.
J. Biol. Chem. 258, 4556-4564, 1983
A:Title: The 5' noncoding and flanking regions of the avian very low density apolipoprotein
A:Reference number: A05078; MUID:83161037; PMID:6187737
A:Accession: A05078
A:Molecule type: DNA
A:Residues: 1-28 <HAC>
A:Cross-references: GB:V00381; NID:G63038; PIDN:CAA23680.1; PID:G63039
R:Rosen, A.M.; Geller, D.M.
Biochem. Biophys. Res. Commun. 78, 1060-1066, 1977
A:Title: Chicken microsomal albumin: amino terminal sequence of chicken proalbumin.
A:Reference number: A13451; MUID:78019943; PMID:9111327
A:Accession: A13451
A:Molecule type: protein
A:Residues: 19-23 'M', 25-30 <ROS>
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
mones (weak bonds with these hormones promote their transfer across the membranes), the
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-26/Domain: propeptide #status predicted <PRO>
F:27-613/Product: serum albumin #status predicted <MAT>
F:32-206/Domain: serum albumin repeat homology <SA1>
F:225-398/Domain: serum albumin repeat homology <SA2>
F:417-596/Domain: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-89,102-118,117-128,152-197,196-205,228-274,273-281,293-307,306-317,344-389,388-397
Query Match 50.2%; Score 1557.5; DB 1; Length 615;
Best Local Similarity 46.7%; Pred. No. 1.6e-95;
Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;
QY 3 HKSEVARRPKDILGSENFALVLIAPQYLQCPEDHVKLVNEVTEFAKTCVADESAENC 62
DB 30 HKSEIARRNYDLKEETFRKAVAMITFAQYLQRCSEGLSKLVKDVVDLQACKVANEDAPEC 89
QY 63 DKSLHTLFGDKLCTVATLRETYGEMADCCAKOBERNECFLOHKDDNPNLRLVPRVDM 121
DB 90 SKPLPSIILDEICQVKLRDYSYGAVADCCSKADPERNECFLSFKVQSQDFVQYQRPASD 149
QY 122 VMCTAFHNEETFFLKKLYEYIARRHPYFYAPPELLFFAKRYKAAFTCCQADKAAALLPK 181
DB 150 VTCQEQDNVRSFGLHFTYSVARRHPFLYAPAILSFVDFEHALQSCCKESDVGACLDTK 209
QY 182 LDELEDEGKASAKORLKCASIQKFERAFKAVARLSORFPKAEVSKLAVTDLPKV 241
DB 210 ETVMEKAKGVSKVQYFCGLTKQFGDRVFOARQLIYLSQKYPKAPFSEVSKFVDSIGV 269
QY 242 HTECCGGDLLECADRADLAKYICENQDSISSKLKECCFKPLLEKSHCIAEVENDEMPAD 301
DB 270 HKECCGDMVECDMDWARMMGNLCSQQDVFSQKIKDCEKPIVERSQCIIMEAFDEKPAD 329
QY 302 LPSLAADPVESKDVCKNYAEAKDVLGFLMFLYEYARRHPDYSVLLLLAKTYETLEKCC 361
```


Best Local Similarity 75.0%; Pred. No. 28-153;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHKEVAHFKDGLGEENFKALVLIAPQYLOQCPFEDHVKLVNVEVTEFAKTCVADSEAE 60
DB 25 DTHKSEIAHFKDGLGEENFKALVLIAPQYLOQCPFEDHVKLVNVEVTEFAKTCVADSEAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREV 120
DB 85 GCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREV 143
QY 121 DVMCTAFHNEETFLKLYIEARHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
DB 144 DTLCAEFKADKKFWGKLYIEARHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 203
QY 181 KIDELADEGKASSAKORLKASQKFGGERAFKAWAVARLSQRPKAFPAEVSCLVTDLT 240
DB 204 KIDAMEKVLASSARQRLCASIQKFGGERALKAWAVARLSQRPKAFPAEVSCLVTDLT 263
QY 241 VHTCECHGDLLECCADRDADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVNDMPA 300
DB 264 VHKCECHGDLLECCADRDADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVNDMPA 323
QY 301 DLPSLAADFVESKDYCKNKAEDKVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
DB 324 NLPLTADFAEDKEVKCKYQKADKVLGSLFYESSRRHPDYAVSVLLRLAKTYETTLK 383
QY 361 CAADPHCYAKYVFEFKPLVEEPQNLKQNCLELGEYKFNALLVRYTKKQPVST 420
DB 384 CAKEDPHCYAKYVFEFKPLVEEPQNLKQNCLELGEYKFNALLVRYTKKQPVST 443
QY 421 PTLVEISRLGKVGSKCKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCTES 480
DB 444 PTLVEISRLGKVGSKCKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCTES 503
QY 481 LVNRRPCFSALEVDVETVPKFNATFTTHADICTLSEKERQIKKOTALVELVHKPKAT 540
DB 504 LVNRRPCFSALEVDVETVPKFNATFTTHADICTLSEKERQIKKOTALVELVHKPKAT 563
QY 541 KEQLKAWMDFAAFVEKCKKADDEKTCFAEKGKLVAAASQAL 583
DB 564 DEQKUTVNEFVAFVDDKCCADDEKGCFLVLEKPLVASTQAL 606

RESULT 7
ABRTS
serum albumin precursor - rat
N;Alternate names: preproalbumin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-May-1979 #sequence revision 31-May-1979 #text change 22-Jun-1999
C;Accession: A93872; A92211; A91946; A91940; C45800; 157621; A03233
R;Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A;Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A;Reference number: A93872; MUID:81223722; PMID:7017712
A;Accession: A93872
A;Molecule type: mRNA
A;Residues: 1-608 <SAR>
A;Cross-references: GB:J00698; NID:G55627; PIDN:CAA24532.1; PID:G55628
R;Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A;Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
A;Reference number: A92211; MUID:77243657; PMID:893447
A;Note: cleavages during protein maturation
A;Accession: A92211
A;Molecule type: protein
A;Residues: 1-38 <STR>
R;Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A;Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
A;Reference number: A91946; MUID:78109429; PMID:564345
A;Accession: A91946
A;Molecule type: protein

A;Residues: 25-222 <ISI>
R;Isemura, S.; Ikenaka, T.
J. Biochem. 73, 1183-1196, 1976
A;Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amino
A;Reference number: A91940; MUID:76260153; PMID:956149
A;Accession: A91940
A;Molecule type: protein
A;Residues: 223-288/572-608 <IS2>
A;Note: 262-Leu was also found
R;Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A;Title: Copper(II)-binding ability of human alpha-fetoprotein.
A;Reference number: A90758; MUID:79001617; PMID:80265
A;Contents: annotation; copper binding
R;Caraway, R.E.; Cochran, D.G.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid protease
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: C45800
A;Status: preliminary
A;Molecule type: protein
A;Residues: 166-173 <CAR>
R;Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A;Title: Determinants of rat albumin promoter tissue specificity analyzed by an improve
A;Reference number: 157621; MUID:87286876; PMID:3475566
A;Accession: 157621
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: GB:M16925; NID:G202828; PIDN:AAA40712.1; PID:G554412
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; plasma
F;1-18/Domain: signal sequence #status experimental <SIG>
F;19-24/Domain: propeptide #status experimental <PRO>
F;25-608/Product: serum albumin #status experimental <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>
F;27/Binding site: copper (His) #status experimental
F;77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393.

Query Match 78.2%; Score 2426; DB 1; Length 608;
Best Local Similarity 73.4%; Pred. No. 5.5e-153;
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

QY 1 DAHKEVAHFKDGLGEENFKALVLIAPQYLOQCPFEDHVKLVNVEVTEFAKTCVADSEAE 60
DB 25 EAHKSEIAHFKDGLGEENFKALVLIAPQYLOQCPFEDHVKLVNVEVTEFAKTCVADSEAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLRVREV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPPQFPEA 144
QY 121 DVMCTAFHNEETFLKLYIEARHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 180
DB 145 EAMCTSFQENPTSFGLHYLHEVARRHPYFYAPPELLFFAKRYKAAFTCCQADKAAACLLP 204
QY 181 KIDELADEGKASSAKORLKASQKFGGERAFKAWAVARLSQRPKAFPAEVSCLVTDLT 240
DB 205 KIDAVKEKALVAARVQRWKSSNQFGERAFKAWAVARLSQRPKAFPAEVSCLVTDLT 264
QY 241 VHTCECHGDLLECCADRDADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVNDMPA 300
DB 265 INKCECHGDLLECCADRDADLAKYICENQDSISSKLKECEKPLLEKSHCIAEVNDMPA 324
QY 301 DLPSLAADFVESKDYCKNKAEDKVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDYCKNKAEDKVLGMFLYEVARRHPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHCYAKYVFEFKPLVEEPQNLKQNCLELGEYKFNALLVRYTKKQPVST 420
DB 385 CAGDPPACYGTVLAEFQPLVEEPQNLKQNCLELGEYKFNALLVRYTKKQPVST 444

A;Residues: 1-32 <MAG>
 R;Hsieh, J.C.; Lin, F.P.; Tam, M.F.
 Anal. Biochem. 170, 1-8, 1988
 A;Title: Electrophoretic transfer of an analytical isoelectrofocusing gel
 A;Reference number: A60808; MUID:88267456; PMID:3389500
 A;Accession: B60808
 A;Molecule type: protein
 A;Residues: 25-41 <HSI>
 R;Strawich, E.; Glimcher, M.J.
 Eur. J. Biochem. 191, 47-56, 1990
 A;Title: Toxigenic 'enamelins' identified mainly as serum proteins. Major 'enamelin' is albumin
 A;Reference number: S10780; MUID:90336641; PMID:2379503
 A;Accession: S10780
 A;Molecule type: protein
 A;Residues: 25-41, 'H', '43-57, 59-64 <STR>
 R;Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
 J. Immunol. 143, 1680-1684, 1989
 A;Title: Structures of histamine-releasing peptides formed by the action of acid proteases
 A;Reference number: A45800; MUID:89341406; PMID:2474609
 A;Accession: A45800
 A;Molecule type: protein
 A;Residues: 163-172 <CAR>
 R;Carraway, R.E.; Mitra, S.P.; Cochran, D.E.
 J. Biol. Chem. 262, 5968-5973, 1987
 A;Title: Structure of a biologically active neurotensin-related peptide obtained from pe
 A;Reference number: A26693; MUID:87194805; PMID:2437111
 A;Accession: A26693
 A;Molecule type: protein
 A;Residues: 165-172, 'L', 'CA2>
 R;Reed, R.G.; Putnam, F.W.; Peters Jr., T.
 Biochem. J. 191, 867-869, 1980
 A;Title: Sequence of residues 400-403 of bovine serum albumin.
 A;Reference number: A90309; MUID:82023364; PMID:7283978
 A;Accession: A90309
 A;Molecule type: protein
 A;Residues: 402-433 <RE>
 R;Brown, J.R.
 Fed. Proc. 34, 591, 1975
 A;Title: Structure of bovine serum albumin.
 A;Reference number: A91458
 A;Accession: A91458
 A;Molecule type: protein
 A;Residues: 25-41, 'H', '43-117, 'EQ', 120-179, 181-189, 'E', 191-194, 'A', 215-288, 'E'
 R;Brown, J.R.
 submitted to the Atlas, April 1975
 A;Reference number: A94551
 A;Accession: A94551
 A;Molecule type: protein
 A;Residues: 190-195

 R;Brown, J.R.
 Fed. Proc. 33, 1389, 1974
 A;Reference number: A91457
 A;Contents: annotation; disulfide bonds
 R;Werlen, R.C.; Offord, R.E.; Rose, K.
 Biochem. J. 302, 907-911, 1994
 A;Title: Preparation and characterization of novel substrates of insulin proteinase (EC
 A;Reference number: S55232; MUID:95031935; PMID:7945219
 A;Accession: S55232
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 529-536; 569-572 <WER>
 C;Superfamily: serum albumin; serum albumin repeat homology
 C;Keywords: carrier protein; copper binding; duplication; plasma
 F;1-18/Domain: signal sequence #status experimental <SIG>
 F;19-24/Domain: propeptide #status experimental <PRO>
 F;25-607/Product: serum albumin repeat homology <MAT>
 F;29-201/Domain: serum albumin repeat homology <SA1>
 F;220-393/Domain: serum albumin repeat homology <SA2>
 F;412-591/Domain: serum albumin repeat homology <SA3>
 F;27/Binding site: copper (His) #status predicted
 F;77-86; 99-115; 114-125; 147-192; 191-200; 223-269; 268-276; 288-302; 301-312; 339-384; 383-392; 4

Query Match

78.8%; Score 2446.5; DB 1; Length 607;

Best Local Similarity 75.6%; Pred. No. 2.4e-154;
 Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHSEVAHFKDGLGEEFKALVLIAPQYLOCPEDHVKLVNVEYFPAKT CVADESAAE 60
 DB 25 DTHKSEIAHFKDGLGEEFKALVLIAPQYLOCPEDHVKLVNVEYFPAKT CVADESAAE 84
 QY 61 NCDKSLHTLFGDKLCTVATIREYGEWADCCAKOEPENECFLOHKKDDNPNLRLVRPEV 120
 DB 85 GCEKSLHTLFGDKLCTVATIREYGEWADCCAKOEPENECFLOHKKDDNPNLRLVRPEV 143
 QY 121 DVMCTAFHNDNEFTLKKLYEIAARRHPFYFAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
 DB 144 NTLCEFKADEKKFWGKLYEIAARRHPFYFAPPELLFFAKYKAAFTCCQAADKAAACLLP 203
 QY 181 KLDELREGKASSAKQRLKASLOKGERAFKAWARLRSORPKFAEFVSKLVTDLTK 240
 DB 204 KIETMRKVLASSARQRLKASLOKGERAFKAWARLRSORPKFAEFVSKLVTDLTK 263
 QY 241 VHTCCCHGDLLEACADRDADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 264 VHTCCCHGDLLEACADRDADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 323
 QY 301 DLPSLADFVESKDVCKNYAEAKDVFLGMLYEVARHPDYSVVLRLAKTVEETLEKC 360
 DB 324 NLPLTADFADKDVCKNYAEAKDVFLGMLYEVARHPDYSVVLRLAKTVEETLEKC 383
 QY 361 CAADPHCEYAKVDFEPLVEEPQNLIKONCELFQELGEYKFNALLVRYTKVPQVST 420
 DB 384 CAKDDPHACYSITVFDKHLVDEPNLIKONCELFQELGEYKFNALLVRYTKVPQVST 443
 QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCADYLSVNLQCLVHEKTPVSDRVTKCCTES 480
 DB 444 PTLVEVSRLGKVGSKCKHPEAKRMPCADYLSVNLQCLVHEKTPVSDRVTKCCTES 503
 QY 481 LVNRRPCFSALVEDETVPKFEFNAETTFHADICTLSEKEROIKQKQALVELVHKPKAT 540
 DB 504 LVNRRPCFSALVEDETVPKFEFNAETTFHADICTLSEKEROIKQKQALVELVHKPKAT 563
 QY 541 KEQLKAVMDDFAAFVEKCKKADDETCFAEKGKLVAAASQAL 583
 DB 564 EEQLKTVMENFVAFVCKCAADDETCFAEKGKLVAAASQAL 606

RESULT 6

ASSHS

serum albumin precursor - sheep
 C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C;Accession: S06936
 R;Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
 Nucleic Acids Res. 17, 10495, 1989
 A;Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
 A;Reference number: S06936; MUID:90098888; PMID:2602160
 A;Accession: S06936

A;Molecule type: mRNA

A;Residues: 1-607 <BRO>

A;Cross-references: ENBL.X17055; NID:gl386; PIDN:CAA34903.1; PID:gl387

C;Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper,
 steroid hormones (weak bonds with these hormones promote their transfer across the mem
 C;Superfamily: serum albumin; serum albumin repeat homology

C;Keywords: carrier protein; duplication; metal binding; plasma

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-24/Domain: propeptide #status predicted <PRO>

F;25-607/Product: serum albumin repeat homology <MAT>

F;29-201/Domain: serum albumin repeat homology <SA1>

F;220-393/Domain: serum albumin repeat homology <SA2>

F;412-591/Domain: serum albumin repeat homology <SA3>

F;27/Binding site: copper (His) #status predicted

F;77-86; 99-115; 114-125; 147-192; 191-200; 223-269; 268-276; 288-302; 301-312; 339-384; 383-392; 4

F;263/Binding site: bilirubin (Lys) #status predicted

Query Match

78.4%; Score 2432.5; DB 1; Length 607;

Query Match 84.4%; Score 2620; DB 2; Length 608;
 Best Local Similarity 82.0%; Pred. No. 8.1e-166;
 Matches 478; Conservative 52; Mismatches 53; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYIQOCPFFEDHVKLVNEVTEFAKTCVADSSAE 60
 DB 25 BAHQSEIAHRFNDLGEHFRGLVLIAPAYIQOCPFFEDHVKLVNEVTEFAKTCVADSSAA 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLVPEV 120
 DB 85 NCKSLHLLGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLVPEV 144

QY 121 DVMTAFHNEBTEFLKXLYEYARHPYFAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 DB 145 DAMCTAFHNEBTEFLKXLYEYARHPYFAPELLFFAKRYKAAFTCCQAAADKAACLTP 204

QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEYSLKLVDTLTK 240
 DB 205 KYDALREKVLASSAKERLKASLOKFGGERAFKAWAVARLSORFPKAEFAEYSLKLVDTLAK 264

QY 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 265 IHKECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDELPA 324

QY 301 DPLSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLTK 360
 DB 325 DPLSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLTK 384

QY 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
 DB 385 CATDDPPACYAHVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 444

QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDVLSVNLQCLVLEHKTVPVSDRVTKCCTES 480
 DB 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDVLSVNLQCLVLEHKTVPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTTHADICTLSEKROIKKQATLVELVKKHPKAT 540
 DB 505 LVNRRPCFSALEVDETYVPKFEFNAETFTTHADICTLSEKROIKKQATLVELVKKHPKAT 564

QY 541 KEQLKAVMDDFAAVFEKCKKADDKETCFABEGKLVAAQAAL 583
 DB 565 BEQLKTMGDFGSDVFDKCAAEDEKACFAEFGKLVAAQAAL 607

RESULT 4
 ABBOS
 serum albumin precursor - horse
 C/Species: Equus caballus (domestic horse)
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C/Accession: S34053
 R/Ho, J.X.; Holowachuk, E.W.; Norton, E.J.; Twigg, P.D.; Carter, D.C.
 Eur. J. Biochem. 215, 205-212, 1993
 A/Title: X-ray and primary structure of horse serum albumin (Equus caballus) at 0.27-nm
 A/Reference number: S34053; MUID:93345495; PMID:8344282
 A/Accession: S34053
 A/Molecule type: mRNA
 A/Residues: 1-607 <HOA>
 A/Cross-references: GB:X74045; NID:G399671; PID:CAA52194.1; PID:G399672
 C/Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds copper, teroid hormones (weak bonds with these hormones promote their transfer across the membrane)
 C/Superfamily: serum albumin; serum albumin repeat homology
 C/Keywords: carrier protein; duplication; metal binding; plasma
 F/1-18/Domain: signal sequence #status predicted <SIG>
 F/19-24/Domain: propeptide #status predicted <PRO>
 F/25-607/Product: serum albumin #status predicted <MAT>
 F/25-201/Domain: serum albumin repeat homology <SA1>
 F/220-393/Domain: serum albumin repeat homology <SA2>
 F/412-591/Domain: serum albumin repeat homology <SA3>
 F/21/Binding site: copper (His) #status predicted
 F/77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-392,4
 F/263/Binding site: bilirubin (Lys) #status predicted

Query Match 79.8%; Score 2475.5; DB 1; Length 607;
 Best Local Similarity 76.3%; Pred. No. 2.9e-156;
 Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAYIQOCPFFEDHVKLVNEVTEFAKTCVADSSAE 60
 DB 25 DTHKSTIAHRFNDLGEHFRGLVLIAPAYIQOCPFFEDHVKLVNEVTEFAKTCVADSSAE 84

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLVPEV 120
 DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDPNLPLVPEV 143

QY 121 DVMTAFHNEBTEFLKXLYEYARHPYFAPELLFFAKRYKAAFTCCQAAADKAACLLP 180
 DB 144 DAOCAAFQDDPKFLKXLYEYARHPYFAPELLFFAKRYKAAFTCCQAAADKAACLLP 203

QY 181 KLDELDEGKASSAKORLKASLOKFGGERAFKAWAVARLSORFPKAEFAEYSLKLVDTLTK 240
 DB 204 KLDALKERILLSSAKERLKASLOKFGGERAFKAWAVARLSORFPKAEFAEYSLKLVDTLTK 263

QY 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 DB 264 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDELPS 323

QY 301 DPLSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLTK 360
 DB 324 DPLSLAADFVESKDVCKNYAEAKDVLGMLFYEARHPDYSVLLRLAKTYETTLTK 383

QY 361 CAADAPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
 DB 384 CAADAPHECYAKVDFEFKPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 443

QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDVLSVNLQCLVLEHKTVPVSDRVTKCCTES 480
 DB 444 PTLVEVSRNLGKVGSKCKHPKAPCAEDVLSVNLQCLVLEHKTVPVSDRVTKCCTES 503

QY 481 LVNRRPCFSALEVDETYVPKFEFNAETFTTHADICTLSEKROIKKQATLVELVKKHPKAT 540
 DB 504 LVNRRPCFSALEVDETYVPKFEFNAETFTTHADICTLSEKROIKKQATLVELVKKHPKAT 563

QY 541 KEQLKAVMDDFAAVFEKCKKADDKETCFABEGKLVAAQAAL 583
 DB 564 KEQLKAVMDDFAAVFEKCKKADDKETCFABEGKLVAAQAAL 606

RESULT 5
 ABBOS
 serum albumin precursor [validated] - bovine
 N/Alternate names: 67K protein; preproalbumin
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
 C/Accession: A38885; A36401; A31258; B60808; S10780; D45800; A26693; A90309; A91458; A9
 R/Holowachuk, E.W.; Stolttenborg, J.K.; Reed, R.G.; Peters Jr., T.
 submitted to the EMBL Data Library, August 1991
 A/Description: Bovine serum albumin: cDNA sequence and expression.
 A/Reference number: A38885
 A/Accession: A38885
 A/Molecule type: mRNA
 A/Residues: 1-607 <HOL>
 A/Cross-references: EMBL:J73215
 R/Hitayama, K.; Akashi, S.; Furuya, M.; Fukuhara, K.
 Biochem. Biophys. Res. Commun. 173, 639-646, 1990
 A/Title: Rapid confirmation and revision of the primary structure of bovine serum albumin
 A/Reference number: A36401; MUID:91083649; PMID:2260975
 A/Accession: A36401
 A/Molecule type: protein
 A/Residues: 25-41, 'H', '43-189, 'E', '191-213, 'T', '215-323, 'D', '325-393, 'TS', '396-607 <HTR>
 R/MacGillivray, R.T.A.; Chung, D.W.; Davie, E.W.
 Eur. J. Biochem. 98, 477-485, 1979
 A/Title: Bioynthesis of bovine plasma proteins in a cell-free system.
 A/Reference number: A31258; MUID:80024278; PMID:488109
 A/Accession: A31258
 A/Molecule type: protein

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants
R;Menaya J.; Parrilla R.; Ayuso, M.S.
submitted to the EMBL Data Library, March 1995
A;Reference number: G08292
A;Accession: G01747
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-120,'G',122-455 <MEN>
A;Cross-references: EMBL:U22961; NID:G763428; PID:AAA64922.1; PID:G763431
P;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
Biochem. J. 308, 321-325, 1995
A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast Kex2
A;Reference number: S55314; MUID:95275251; PMID:7755581
A;Accession: S55314
A;Molecule type: protein
A;Residues: 19-27 <LED>
P;Meloun, B.; Moravsek, L.; Kostka, V.
FEBS Lett. 56, 134-137, 1975
A;Title: Complete amino acid sequence of human serum albumin.
A;Reference number: A91420; MUID:76187907; PMID:1225573
A;Accession: A91420
A;Molecule type: protein
A;Residues: 25-117,'EQ',120-154,'Q',156-193,'E',195-387,'H',389-390,'Y',392-393,'A',395-
P;Roehr, U.; Spitzler, G.; Tripler, D.
Justus Liebig's Ann. Chem. 9, 881-884, 1988
A;Title: Isolation and structure elucidation of middle-molecular weight peptides from u
A;Reference number: S06422
A;Note: this paper is in German, with an English abstract
A;Accession: S06422
A;Molecule type: protein
A;Residues: 25-48 <ROE>
P;Finch, J.W.; Crouch, R.K.; Knaop, D.R.; Schey, K.L.
Arch. Biochem. Biophys. 305, 595-599, 1993
A;Title: Mass spectrometric identification of modifications to human serum albumin treat
A;Reference number: S36882; MUID:93384321; PMID:8733198
A;Accession: S36882
A;Molecule type: protein
A;Residues: 45-67,'141-160',311-337,'469-490',570-581 <FIN>
P;Kausler, E.; Spitzler, G.
Biochem. Biophys. Res. Commun. 136, 849-855, 1991
A;Title: Bruchstuecke aus Albumin und beta(2)-Mikroglobulin - Bestandteile der Mittelmol
A;Reference number: S17599; MUID:92126241; PMID:1772598
A;Accession: S17599
A;Molecule type: protein
A;Residues: 25-54,'354-357',431-447 <KAU>
A;Note: 49-Leu was also found
P;Caraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1580-1584, 1989
A;Title: Structures of histamine-releasing peptides formed by the action of acid proteas
A;Reference number: A45800; MUID:89341406; PMID:2474609
A;Accession: A45800
A;Molecule type: protein
A;Residues: 166-173 <CAR>
P;Mogard, M.H.; Kobayashi, R.; Chen, C.F.; Lee, T.D.; Reeve Jr., J.R.; Shively, J.E.; Wa
Biochem. Biophys. Res. Commun. 136, 993-998, 1986
A;Title: The amino acid sequence of kinetins, a novel peptide isolated from pepsin-tre
A;Reference number: A03239; MUID:86242180; PMID:3087352
A;Accession: A03239
A;Molecule type: protein
A;Residues: 166-173,'L', <MOG>
P;Galliano, M.; Minchiotti, L.; Porta, F.; Rossi, A.; Ferri, G.; Madison, J.; Watkins, S
Proc. Natl. Acad. Sci. U.S.A. 87, 8721-8725, 1990
A;Title: Mutations in genetic variants of human serum albumin found in Italy.
A;Reference number: A38255; MUID:91062352; PMID:2247440
A;Accession: A38255
A;Molecule type: protein
A;Residues: 76-111 <GAL1>
A;Accession: B38255
A;Molecule type: protein
A;Residues: 82-105,'K',107-110 <GAL2>
A;Note: this variant is designated albumin Vibo Valentia
A;Accession: A38255
A;Molecule type: protein

A;Residues: 76-83,'K',85-106 <GAL3>
A;Note: this variant is designated albumin Torino
P;Minchiotti, L.; Galliano, M.; Zapponi, M.C.; Tenni, R.
Eur. J. Biochem. 214, 437-444, 1993
A;Title: The structural characterization and bilirubin-binding properties of albumin He
A;Reference number: S33298; MUID:93292504; PMID:8513793
A;Accession: S33298
A;Molecule type: protein
A;Residues: 255-263,'E',265-281 <MIN1>
A;Note: this variant is designated albumin Herborn
P;Minchiotti, L.; Galliano, M.; Stoppini, M.; Ferri, G.; Crespeau, H.; Rochu, D.; Porta
Biochim. Biophys. Acta 1119, 232-238, 1992
A;Title: Two albumins with identical electrophoretic mobility are produced by diffe
A;Reference number: S21078; MUID:92190239; PMID:1347703
A;Accession: S21078
A;Molecule type: protein
A;Residues: 354-356,'K',358-378 <MIN2>
A;Note: this variant is designated albumin Sondrio; another variant Paris-2 is reported
P;He, X.M.; Carter, D.C.
Nature 358, 209-215, 1992
A;Title: Atomic structure and chemistry of human serum albumin.
A;Reference number: A46756; MUID:92334427; PMID:1630489
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
P;Brown, J.R.; Shockley, P.; Behrens, P.Q.
in The Chemistry and Physiology of the Human Plasma Proteins, Bing, D.H., ed., pp.23-40
A;Reference number: A94442
A;Contents: annotation; three-dimensional structure and disulfide bonds
P;Saber, M.A.; Stockbauer, P.; Moravsek, L.; Meloun, B.
Collect. Czech. Chem. Commun. 42, 564-579, 1977
A;Title: Disulfide bonds in human serum albumin.
A;Reference number: A90930
A;Contents: annotation; disulfide bonds
P;Jacobsen, C.
Biochem. J. 171, 453-459, 1978
A;Title: Lysine residue 240 of human serum albumin is involved in high-affinity binding
A;Reference number: A90299; MUID:78186630; PMID:656055
A;Contents: annotation; bilirubin-binding site
P;Peters, T.; Reed, R.G.
in Albumin: Structure, Biosynthesis, Function, Peters, J., and Sjolholm, I., eds., 11-20
A;Title: Serum albumin: conformation and active sites.
A;Reference number: A94408
A;Contents: annotation; binding sites
P;Harper, M.E.; Dugalczyk, A.
Am. J. Hum. Genet. 35, 565-572, 1983
A;Title: Linkage of the evolutionarily-related serum albumin and alpha-fetoprotein gene
A;Reference number: A90028; MUID:8327982; PMID:6192711
A;Contents: annotation; gene position
P;Walker, J.E.
FEBS Lett. 66, 173-175, 1976
A;Title: Lysine residue 199 of human serum albumin is modified by acetylsalicylic acid
A;Reference number: A46755; MUID:76257808; PMID:955075
A;Contents: annotation
P;Bohney, J.P.; Fonda, M.L.; Feldhoff, R.C.
FEBS Lett. 298, 266-268, 1992
A;Title: Identification of Lys(190) as the primary binding site for pyridoxal 5'-phosph
A;Reference number: A56234; MUID:92183981; PMID:1544460
A;Contents: annotation
A;Note: the nonenzymatic binding of pyridoxal phosphate to lysine-214 is described; in
case activity
C;Comment: Serum albumin, a predominant protein in the plasma of adults, is synthesized
lirubin, protoporphyrin, long-chain fatty acids, prostaglandins, steroid hormones (weak
C;Comment: A large number of variants of human serum albumin have been described.
C;Genetics:
A;Gene: GDB:ALB
A;Cross-references: GDB:118990; OMIM:103600
A;Map position: 4q11-4q13
C;Superfamily: serum albumin; serum albumin repeat homology
C;Keywords: carrier protein; duplication; metal binding; phosphoprotein; plasma; pyrid
P;1-18/Domain: signal sequence #status predicted <SIG>
P;19-24/Domain: propeptide #status experimental <PRO>
P;25-609/Product: serum albumin #status experimental <MPT>
P;25-202/Domain: serum albumin repeat homology <SA1>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 11:37:59 ; Search time 71.3019 Seconds
(without alignments)
789.208 Million cell updates/sec

Title: US-09-832-929-18
Perfect score: 3103
Sequence: 1 DAHSEVAHRFKDLGEENFK.....TCFAEKGKLVASQAALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : PIR 78:*

- 1: pir1:*
- 2: pir2:*
- 3: pir3:*
- 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3103	100.0	609	1 ABHUS	serum albumin prec
2	2942	94.8	600	2 A47391	serum albumin prec
3	2620	84.4	608	2 S57632	serum albumin prec
4	2475.5	79.8	607	1 ABHOS	serum albumin prec
5	2446.5	78.8	607	1 ABBSH	serum albumin prec
6	2432.5	78.4	607	1 ABBSH	serum albumin prec
7	2426	78.2	608	1 ABRTS	serum albumin prec
8	2411.5	77.7	605	1 ABPGS	serum albumin prec
9	2387	76.9	609	2 JCS938	albumin - Mongolia
10	1861	60.0	453	2 A05139	serum albumin - mo
11	1557.5	50.2	615	1 ABCHS	serum albumin prec
12	1253.5	40.4	609	2 JCA258	alpha-fetoprotein
13	1249.5	40.3	609	1 PPHU	alpha-fetoprotein
14	1242.5	40.0	609	1 PPGU	alpha-fetoprotein
15	1205	38.8	607	1 ABXL72	74K albumin prec
16	1181.5	38.1	265	2 I46986	albumin - dog (fra
17	1175.5	37.9	608	1 ABXL68	68K serum albumin
18	1084	34.9	605	1 PPM5	alpha-fetoprotein
19	1057	34.4	611	1 PPRT	alpha-fetoprotein
20	1055	34.0	599	1 A54906	afanin precursor -
21	928.5	29.9	614	2 S59517	serum albumin prec
22	928	29.9	608	2 A53195	afanin precursor -
23	747.5	24.1	608	1 ABONS1	serum albumin 1 pr
24	742.5	23.9	608	1 ABONS2	serum albumin 2 pr
25	699	22.5	382	2 A37253	serum albumin - bu
26	440.5	14.2	1423	1 S27941	serum albumin - se
27	386	12.4	474	1 VYHUD	vitamin D-binding
28	385	12.4	476	1 VYRID	vitamin D-binding
29	372	12.0	472	1 A35327	vitamin D-binding

ALIGNMENTS

RESULT 1

ABHUS

serum albumin precursor [validated] - human

N;Alternate names: preproalbumin

C;Species: Homo sapiens (man)

C;Date: 29-Jul-1981 #sequence, revision 31-Jan-1997 #text change 17-Mar-2000

C;Accession: A93743; A93936; J19427; I59286; I59313; G01747; S55314; A91420; S06422; S3

R;Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, A.E.; Houch, C.M.; Najarian, R.C.; Seebu

Nucleic Acids Res. 9, 6103-6114, 1981

A;Title: The sequence of human serum albumin cDNA and its expression in Escherichia col

A;Reference number: A93743; MUID:82081882; PMID:6171778

A;Accession: A93743

A;Molecule type: mRNA

A;Residues: 1-419, 'K', 421-609 <LAW>

A;Cross-references: EMBL:V00495; GB:J00078; GB:L00132; GB:L00133; NID:G28591; PIDN:CAA2

R;Dugaiczky, A.; Law, S.W.; Dennison, O.E.

Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982

A;Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA.

A;Reference number: A93936; MUID:82105994; PMID:6275391

A;Accession: A93936

A;Molecule type: mRNA

A;Residues: 1-120, 'G', 122-609 <DUG>

A;Cross-references: EMBL:V00494; NID:G28589; PIDN:CAA23753.1; PID:G28590

R;Urano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.

J. Biol. Chem. 261, 3244-3251, 1986

A;Title: The human albumin gene. Characterization of the 5' and 3' flanking regions and

A;Reference number: I39427; MUID:86140099; PMID:2419329

A;Accession: I39427

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-26 <URA>

A;Cross-references: GB:M13075; NID:G178330; PIDN:AAAS1688.1; PID:G553173

R;Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.

Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994

A;Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian family

A;Reference number: I59286; MUID:94181575; PMID:8134387

A;Accession: I59286

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 282-290, 'KSRFDLQ' <WAT>

A;Cross-references: GB:S69192; NID:G546032; PIDN:AAB30282.1; PID:G546033

A;Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia

R;Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putnam,

Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994

A;Title: Genetic variants of human serum albumin in Italy: point mutants and a carboxyl

A;Reference number: I59313; MUID:94294404; PMID:8022807

A;Accession: I59313

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 589-590, 'ALPRVKNLLQVKLP' <WAD>

A;Cross-references: GB:S70799; NID:G547231; PIDN:AAB31177.1; PID:G547232


```
QY 301 DLPSLAADFVSKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVSKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
Db 361 CAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 14
US-10-153-604A-5
; Sequence 5, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-5

Query Match 100.0%; Score 2068; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-130;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDGLGENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADES 60
Db 1 DAHKEVAHRFKDGLGENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADES 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREP 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREP 120
QY 121 DVNCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
Db 121 DVNCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 300
Db 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 300
QY 301 DLPSLAADFVSKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVSKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
Db 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
```

```
RESULT 15
US-10-319-263-1
; Sequence 1, Application US/10319263
; Publication No. US20030180820A1
; GENERAL INFORMATION:
; APPLICANT: Bar-Or M.D., David
; APPLICANT: Iau Ph.D., Edward
; APPLICANT: Winkler M.D., James V.
; TITLE OF INVENTION: Tests for the Rapid Evaluation of Ischemic States and
; TITLE OF INVENTION: Kits
; FILE REFERENCE: ISC007
```

```
; CURRENT APPLICATION NUMBER: US/10/319,263
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 60/115,392
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: 60/102,738
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,926
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 09/165,581
; PRIOR FILING DATE: 1998-10-02
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-319-263-1

Query Match 100.0%; Score 2068; DB 14; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKEVAHRFKDGLGENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADES 60
Db 1 DAHKEVAHRFKDGLGENFKALVLIAPAQYLQCCPFEDHVKLVNVEVTEFAKTCVADES 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREP 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVREP 120
QY 121 DVNCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
Db 121 DVNCTAFHDNEETFLKKLYEYIARRHPYFAPPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSORFPKAEFAEYVKLVTDLT 240
QY 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 300
Db 241 VHTECHGDLLECCADRADLAKYICENQDSISSKLKECCEKPLLEKSHCIAEVENDEMP 300
QY 301 DLPSLAADFVSKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
Db 301 DLPSLAADFVSKDVCKNYAEAKDVFGLGMFLYEYARRHPDYSVLLRLAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
Db 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
```

Search completed: April 19, 2004, 12:54:57
Job time : 163.947 secs

```

; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 2068; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
DB 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYIYARRHPDYSVVLRLRAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYIYARRHPDYSVVLRLRAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEBPQNL 387
DB 361 CAADPHECYAKVDFEFKPLVEBPQNL 387

RESULT 13
US-10-153-064-5
; Sequence 5, Application US/10153064
; Publication No. US20020142814A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: P7556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 2002-05-24
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-5

Query Match      100.0%; Score 2068; DB 13; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
DB 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300

; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-31

Query Match      100.0%; Score 2068; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHKSEVAHRFKDLGEENFKALVLIAPAOYLQCCPPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKLYEIAARRHPYFYAPPELLFFAKYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
DB 181 KLDELDRDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFAEVSCLVTDLT 240
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300
DB 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLECCCKPILLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYIYARRHPDYSVVLRLRAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYIYARRHPDYSVVLRLRAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEBPQNL 387
DB 361 CAADPHECYAKVDFEFKPLVEBPQNL 387

RESULT 12
US-10-433-108-34
; Sequence 34, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-433-108-34

Query Match      100.0%; Score 2068; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVYARRHPDYSVLLLRLLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVYARRHPDYSVLLLRLLAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
DB 361 CAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 9

US-09-833-245-18
; Sequence 18, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-833-245-18

Query Match 100.0%; Score 2068; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-130;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGLENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGLENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVYARRHPDYSVLLLRLLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVYARRHPDYSVLLLRLLAKTYETTLK 360

QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
DB 361 CAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 10

US-10-424-999-11
; Sequence 11, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 11
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-11

Query Match 100.0%; Score 2068; DB 12; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGLENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHRFKDLGLENFKALVLIAPQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAKRYKAAFTCCQAAADKAAACLLP 180
QY 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
DB 181 KLDELDRDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQFFPKAEFAVSKLVTDLTK 240
QY 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300
DB 241 VHTTECHGDLLECCADDRADLAKYICENQDSISSKKECCCKPILLESKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVYARRHPDYSVLLLRLLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGMFLYEVYARRHPDYSVLLLRLLAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPQNL 387
DB 361 CAADPHECYAKVDFEFKPLVEEPQNL 387

RESULT 11

US-10-425-000-31
; Sequence 31, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit

; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (BLYS)
; FILE REFERENCE: DXX-018.1 PCT; DXX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-932-322-445

Query Match 100.0%; Score 2068; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGGENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADSSAE 60
DB 1 DAHSEVAHFRKDLGGENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADSSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
QY 241 VHTECCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
DB 241 VHTECCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLEKC 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLEKC 360
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387
DB 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387

RESULT 7

US-09-832-501-18
; Sequence 18, Application US/09832501
; Publication No. US20030199043A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J.
; APPLICANT: Sleep, Darrell
; APPLICANT: Turner, Andrew J.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF542
; CURRENT APPLICATION NUMBER: US/09/832,501
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-832-501-18

Query Match 100.0%; Score 2068; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGGENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADSSAE 60
DB 1 DAHSEVAHFRKDLGGENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADSSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
DB 121 DVMTAFHDNEETFLKKYLYEIAARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180
QY 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKASIQKGERAFKAWAVARLSQRPKAEFAEVSCLVTDLT 240
QY 241 VHTECCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
DB 241 VHTECCHGDLLECCADRADLAKYICENQDSISSKLKECCPKLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLEKC 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETTLEKC 360
QY 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387
DB 361 CAADPHCEYAKVDFEFKPLVEEPQNL 387

RESULT 8

US-09-833-118-18
; Sequence 18, Application US/09833118
; Publication No. US20030219875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF544
; CURRENT APPLICATION NUMBER: US/09/833,118
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-118-18

Query Match 100.0%; Score 2068; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGGENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADSSAE 60
DB 1 DAHSEVAHFRKDLGGENFKALVLIAPAFYLOQCPEDHVKLVNEVTEFAKTCVADSSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

QY 121 DVMCTAFHNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKCASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETLEKC 360
QY 361 CAADPHECYAKVDFDEKPLVEEPQNL 387
DB 361 CAADPHECYAKVDFDEKPLVEEPQNL 387

RESULT 4

US-09-833-041-18
; Sequence 18, Application US/09833041
; Publication No. US20030125247A1

GENERAL INFORMATION:

; APPLICANT: Rosen, Craig A.
; APPLICANT: Haseltine, William A.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF545
; CURRENT APPLICATION NUMBER: US/09/833,041
; CURRENT FILING DATE: 2001-04-12
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens

US-09-833-041-18

Query Match 100.0%; Score 2068; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEBENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
DB 1 DAHKSEVAHRFKDLGEBENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKCASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETLEKC 360

DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETLEKC 360
QY 361 CAADPHECYAKVDFDEKPLVEEPQNL 387
DB 361 CAADPHECYAKVDFDEKPLVEEPQNL 387

RESULT 5

US-09-833-117-18
; Sequence 18, Application US/09833117
; Publication No. US20030171267A1

GENERAL INFORMATION:

; APPLICANT: Rosen, Craig A.
; APPLICANT: Sadeghi, Homa
; APPLICANT: Prior, Christopher P.
; APPLICANT: Turner, Andrew J.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF543
; CURRENT APPLICATION NUMBER: US/09/833,117
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo Sapiens

US-09-833-117-18

Query Match 100.0%; Score 2068; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGEBENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
DB 1 DAHKSEVAHRFKDLGEBENFKALVLI AFAQYLQCCPFEDHVKLVNEVTEFAKTCVADES 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMCTAFHNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
DB 121 DVMCTAFHNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240
DB 181 KLDELDEGKASSAKORLKCASLQKFGRAKAWAVARLSORFPKAEFAEVSCLVTDLT 240
QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKECCPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETLEKC 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVLLLRLLAKTYETLEKC 360
QY 361 CAADPHECYAKVDFDEKPLVEEPQNL 387
DB 361 CAADPHECYAKVDFDEKPLVEEPQNL 387

RESULT 6

US-09-932-322-445
; Sequence 445, Application US/09932322
; Publication No. US20030194743A1

GENERAL INFORMATION:

; APPLICANT: Dyax Corp.

```

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2
Query Match 100.0%; Score 2068; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPPKAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPPKAEFAEVSKLVTDLTK 240
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
Qy 361 CAADPHECYAKVDFEKPPLVEBPQNL 387
Db 361 CAADPHECYAKVDFEKPPLVEBPQNL 387

RESULT 3
US-09-984-010-26
; Sequence 26; Application US/09984010
; Publication No. US20030104578A1
; GENERAL INFORMATION:
; APPLICANT: Ballance, David James
; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
; AND SERUM ALBUMIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
; STREET: 1300 I Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,010
; FILING DATE: 21-May-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/091,873
; FILING DATE: 25-JUN-1998
; APPLICATION NUMBER: PCT/GB96/03164
; FILING DATE: 19-DEC-1996
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26
Query Match 100.0%; Score 2069; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-929-552-2
Query Match 100.0%; Score 2068; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Qy 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Qy 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPPKAEFAEVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKASLQKFGERAFKAWAVARLSQRPPKAEFAEVSKLVTDLTK 240
Qy 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
Db 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLKECCCKPILLESKSHCIAEVENDEMPA 300
Qy 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
Db 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEVARRHPDYSVLLRLAKTYETTTLEKC 360
Qy 361 CAADPHECYAKVDFEKPPLVEBPQNL 387
Db 361 CAADPHECYAKVDFEKPPLVEBPQNL 387

RESULT 2
US-09-932-613-445
; Sequence 445; Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Beltzer, James P.
; APPLICANT: Potter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: Dlx-025.1 PCT; Dlx-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 445
; LENGTH: 585
; TYPE: PRT
; ORGANISM: HomoSapiens
US-09-932-613-445
Query Match 100.0%; Score 2068; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-190;
Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DAHSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 1 DAHSEVAHRFKDGLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAAE 60
Qy 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 12:00:25 ; Search time 162.947 Seconds
(without alignments)
654.724 Million cell updates/sec

Title: US-09-832-929-18_COPY_1_387

Perfect score: 2068

Sequence: 1 DAKSEVAHFKDLGEENFK.....ECYAKVFDEPKPLVEEPQNL 387

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/FCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2068	100.0	585	9	US-09-929-552-2
2	2068	100.0	585	10	US-09-932-613-445
3	2068	100.0	585	10	US-09-984-010-26
4	2068	100.0	585	10	US-09-833-041-18
5	2068	100.0	585	10	US-09-833-117-18
6	2068	100.0	585	10	US-09-932-322-445
7	2068	100.0	585	10	US-09-832-501-18
8	2068	100.0	585	11	US-09-833-118-18
9	2068	100.0	585	11	US-09-833-245-18
10	2068	100.0	585	12	US-10-424-999-11
11	2068	100.0	585	12	US-10-425-000-31
12	2068	100.0	585	12	US-10-433-108-34
13	2068	100.0	585	13	US-10-153-064-5
14	2068	100.0	585	14	US-10-153-604A-5
15	2068	100.0	585	14	US-10-319-263-1

16	2068	100.0	585	14	US-10-319-263-2	Sequence 2, Appli
17	2068	100.0	585	14	US-10-414-469-1	Sequence 1, Appli
18	2068	100.0	585	14	US-10-414-469-2	Sequence 2, Appli
19	2068	100.0	585	14	US-10-413-831-1	Sequence 1, Appli
20	2068	100.0	585	14	US-10-413-831-2	Sequence 2, Appli
21	2068	100.0	585	15	US-10-413-832-1	Sequence 1, Appli
22	2068	100.0	585	15	US-10-413-832-2	Sequence 2, Appli
23	2068	100.0	585	15	US-10-414-386-1	Sequence 1, Appli
24	2068	100.0	585	15	US-10-414-386-2	Sequence 2, Appli
25	2068	100.0	585	15	US-10-233-675A-11	Sequence 11, Appli
26	2068	100.0	585	15	US-10-462-262-26	Sequence 26, Appli
27	2068	100.0	604	10	US-09-984-010-7	Sequence 7, Appli
28	2068	100.0	609	10	US-09-919-039-370	Sequence 370, App
29	2068	100.0	609	12	US-10-609-346-12	Sequence 12, Appli
30	2068	100.0	609	13	US-10-153-064-7	Sequence 7, Appli
31	2068	100.0	609	14	US-10-153-604A-7	Sequence 7, Appli
32	2068	100.0	609	14	US-10-365-623-23	Sequence 23, Appli
33	2068	100.0	610	9	US-09-984-186-2	Sequence 2, Appli
34	2068	100.0	610	14	US-10-237-687-2	Sequence 2, Appli
35	2068	100.0	610	14	US-10-237-708-2	Sequence 2, Appli
36	2068	100.0	610	14	US-10-237-866-2	Sequence 2, Appli
37	2068	100.0	610	14	US-10-237-871-2	Sequence 2, Appli
38	2068	100.0	610	14	US-10-237-624-2	Sequence 2, Appli
39	2068	100.0	616	12	US-10-433-108-13	Sequence 13, Appli
40	2068	100.0	624	12	US-10-433-108-16	Sequence 16, Appli
41	2068	100.0	631	12	US-10-433-108-14	Sequence 14, Appli
42	2068	100.0	640	12	US-10-433-108-15	Sequence 15, Appli
43	2068	100.0	640	12	US-10-433-108-17	Sequence 17, Appli
44	2068	100.0	651	13	US-10-153-064-133	Sequence 133, App
45	2068	100.0	651	14	US-10-153-604A-133	Sequence 133, App

ALIGNMENTS

RESULT 1

US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MERI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids